译者的话

道格拉斯·R·霍夫施塔特是美国当代计算机教授。哲学并非他的本行。而他的合作者丹尼尔·C·丹尼特则是哲学教授。1980年，他们在加利福尼亚州帕洛阿尔托的行为科学高级研究中心进行了几次谈话，萌生了编著本书的念头。当时丹尼特正在研究人工智能和哲学。他们两人相辅相成，合作编著了本书，配合得十分默契。

当代认知科学是一门研究领域涉及极广的科学，它汇集了心理学、人工智能、神经科学等各大科学的研究成果。霍夫施塔特为了更全面，更完整，也是为了更谨慎地向读者呈现一幅近二三十年来欧美认知科学的研究概况，在本书中扮演了一个批注人的角色。他从卷帙浩繁的认知科学文献中选择了二十多种富有代表性的书刊，又从其中摘选了一部分富有代表性的章节，然后在每篇末尾加上自己以“反思”为题的评论，汇编成了本书。

哲学往往被不少人描绘成玄而又玄的东西，认知科学更其如此。关于“心灵”、“自我”等这样一些认知科学的中心概念千百年来一直被许多人视作难以求解的奥秘，有时甚至还被染上神秘的色彩，使普通人感到高深莫测。

本书的作者们一反这个传统，都试图从各自不

同的角度对有关心灵、自我等问题作出深入浅出、别出心裁的阐述和讨论，例如他们提的下述问题就是很容易为普通人所理解的：

"我是谁？如果我就是你，那会怎么样？"

"如果你的父母从不曾相遇并相爱，那还会不会有你呢？你会不会是别人家的孩子呢？"

"你若没有脑袋，那会怎么样呢？"

"如果莎士比亚的剧本不是他本人写的，而是由一个同名同姓的人写的，那又会怎么样呢？"

从柏拉图的时代起，哲学家们就一直对像心灵或自我这样的东西是否真正存在的本体论问题进行过苦苦思索。有人提出，从解剖学的角度看，人体内不可能存在像心灵或自我这样的东西；可也有人反驳说，自然界存在着声音这样的东西吗？如果存在，它的本质又是什么？如果这种东西能存在，心灵和自我又何尝不能存在呢？为了深入研究心灵问题，哲学家们曾提出各种各样大胆的假设。英国哲学家约翰·洛克早在1690年就提过这样一个问题：如果一个王子的心灵携带着他的记忆进入了一个修鞋匠的身体，那会发生什么样的情况呢？现代还有些人假设并研究了像大脑移植、人格分裂、人格合并以及人的复制这样的问题。人们的这些努力实质上

反映了人对自我、心灵这些问题始终存在着重重疑虑，总是想设法通过理论的和实践的努力来消除它们，达到真正认识自我的目的。

法国哲学家笛卡儿在1641年曾设想过一个十分著名的思维实验。他这样问自己："我怎么知道自己没有在受一个魔鬼的欺骗呢？这个骗术通天的魔鬼想引诱我去相信存在着外部世界（和我自己的身体）。"也许这位说过"我思故我在"这句名言的笛卡儿认为除了魔鬼以外，唯一存在的东西就是他那不朽的灵魂了。这种怀疑主义的论点至今还在流行。随着当代医学科学的进步，这个古老的思维实验又有了现代化的说法：我怎么知道某些心狠手辣的医生没有趁我熟睡之际切除了我的大脑？这些居心叵测的医生把我可怜的大脑贮存在一个生命维持系统内，用各种各样假的刺激来戏弄、欺骗我的大脑。我的大脑难道不就是我吗？那么，我又在哪里呢？我的心灵同我的大脑又是什么关系呢？

随着当代计算机科学的不断发展和人工智能研究的进一步深入，人们开始以一种新的观点来审视认知科学中的一些传统问题。有人曾提出，大脑就好比计算机硬件，而心灵则好比软件。这也正是本书的某些篇章重点强调并专门讨论的一个引人瞩目的

目次

导　论…………………………………………… 1

1　博尔赫斯和我………………………………… 17
2　无头有感……………………………………… 21
3　心灵的再发现………………………………… 35
4　计算机和智能………………………………… 53
5　图林测试：咖啡馆里的对话………………… 72
6　黄粱公主……………………………………106
7　猩猩玛塔的灵魂……………………………111
8　三号兽的心灵………………………………122
9　精神…………………………………………131
10　自私的基因和自私的米姆…………………137
11　前奏曲……蚂蚁赋格曲……………………165
12　大脑的故事…………………………………222
13　我在哪里？…………………………………238

14 我曾在哪里？……………………………258
15 异体并不排斥……………………………271
16 软件……………………………………286
17 宇宙之谜及其谜解………………………308
18 第七次远足或特鲁尔的徒然自我完善……325
19 心灵、大脑和程序………………………336

导论

你看见月亮从东方升起。你看见月亮从西方升起。你看见两轮月亮穿过漆黑、寒冷的天空，迎面相遇，迅又擦肩而过，各自东西。你在火星上，离家千万里，披着地球上制造的脆弱的薄膜，抵御那火星上红色沙漠刺骨的干寒的伤害，你虽护身有术，但却一筹莫展，因为你的宇宙飞船坏了，无法修复。你将再也不能回到你离开的地球、家庭和故乡，回到朋友的身边。

但是，也许还有希望。在损坏的飞船的通讯舱内，你发现有一台马克4型的长途传真运输机和使用指令。如果你拨亮开关，把信号对准地球上的接收器，然后跨入运输舱，长途运输机就会很快毫无痛苦地把你的身体分解开来，制成一幅由分子组成的蓝图，以备送回地球；而在地球上，接收器（其贮存库内贮满了必要的原子）一收到发来的信号指令，几乎顷刻间就复制出了——你！你从火星上以光速被送返地球，送到你亲人的怀抱中，他们马上将如痴如醉地聆听你讲的火星奇遇。

对破损的宇宙飞船望了最后一眼之后，你觉得长途传真运输机是唯一的一线希望。你完好无缺，打开发射机，拨亮开关，钻进运输舱。5，4，3，2，1，发射！你打开面前的门，钻出接受舱，走进地球上阳光灿烂的故土。你回到亲人中间，一点没

变，只是由于从火星到地球的长途传真而略感疲倦。这次从火星上死里逃生值得庆贺，你的亲友为此欢聚一堂，而你却发现，自从上次一别，每一个人都变化甚大。毕竟是快三年了，你也老多了。看一眼你的女儿萨拉，现在该有八岁半了吧。你不禁思忖："这难道就是那个爱坐在我怀里的小姑娘吗？"你回忆起来了，当然是的，尽管你得承认，你认出了她并非全凭记忆。她比以前高多了，看上去也大多了，而且懂事多了。事实上，上次离别时，她现在身上的大部分细胞都还不存在。但是，尽管年龄的增长和变化，尽管细胞新陈代谢，她仍然是三年前你吻别的那个小女孩。

接着，你心乱如麻，思绪万千："我难道真的是三年前吻别这个小女孩的那个人吗？我是这个八岁孩子的母亲呢，还是实际上成了另一个新人，一个刚刚活了几个小时的新人，尽管我还有着对往日和往年的记忆，甚或明显的记忆。这个小女孩的母亲最近葬身火星了吗？在马克4型长途运输机的舱内被分解、被毁灭了吗？

"我死在火星上了吗？不，我肯定没有死在火星上，因为我现在活在地球上。然而，也许有个人死在火星上了——那是萨拉的母亲。这样，我就不是萨拉的母亲了。可我肯定是她的母亲！钻进长途传真运输机就是为了要回到亲人的中间！但是，我一直很健忘，也许我从没有去过火星上的那个运输舱，如果确有其事，也许那是另外一个人。那台地狱般的机器是台长途传真运输机——一种交通工具呢，还是正如商标上写的那样，是一种杀人的双生子制造机？萨拉的母亲是不是从那台机器里幸存下来了呢？她以为会的。她是怀着希望而不是抱着一死了之的想法走进运输舱的，她的行为是为了他人着想，明确地说，她这样做是为了爱萨拉和保护萨拉；也是为了她自己，她想摆脱困

境，化险为夷。或者说看上去是这样的。我怎么会知道看上去是那样的呢？因为我曾在那里，因为我曾是考虑过这些事情的萨拉的母亲，因为我现在是萨拉的母亲，或者说看上去是这样的。”

打这以后，你的情绪大起大落，欢乐和宽慰的心情被恼人的怀疑和心灵探索的乌云所笼罩。心灵的探索。你也许会这样想，不该与萨拉一起兴高采烈地以为她的母亲回来了。你觉得自己有点像一个骗子，要是有一天萨拉明白火星上的真相，她会怎么想呢？还记得吗？当她明白圣诞老人是怎么回事的时候，显得多么的疑惑和痛苦。她自己的妈妈怎么能一直欺骗她？

于是当你捧起这本《心我论》时，不仅仅是出于理智上随意的好奇心，因为本书明言将送你踏上寻找自我和心灵的长途旅程，从本书中你将得知，你是什么，你是谁。

读到此处，你掩卷深思：

我现在正在读本书第3页。我还活着，我还清醒；我的双眼看得清书中的字；我看见我双手捧书。我有双手。我怎么知道这就是我的手呢？这问题真蠢。因为它们与我的胳膊、我的身体联在一起。我怎么知道这就是我的身体呢？因为它受我控制。那么，我拥有我的身体吗？从某种意义上说，确实如此。只要我不去伤害别人，我尽可以随心所欲。这甚至还是一种法定的财产，因为只要我还活着，我就不能合法地将我的身体卖给任何人，但一旦我的身体死了，我就能够合法地将我身体的所有权出卖，比如说，卖给一所医学院。

如果我拥有这个身体，那么我就认为我不只是一个身体。要是我说“我拥有我的身体”，我的意思不是说“这个身体拥有它自己”，也许这句话并没有什么意思。否则，没被旁人拥有的万物是否拥有了自身？月亮是属于万人，还是不属于任何人，还是

就属于它自己？谁能成为万物的拥有者呢？我就能。我的身体就是我拥有的一份东西。不管怎么样，我和我的身体既紧密相联又互有不同。我是控制者，我的身体是被控制者。大部分情况下是这样。

接着，《心我论》问你：如果那样的话，你是不是想调换一个更强壮、更美丽、更易控制的身体？

你以为这不可能。

但是，本书坚持认为，这完全能够想象，因此原则上是可能的。

你怀疑本书是不是含有灵魂的再生和轮回的思想。本书预见到了这种疑问并承认，尽管再生这个想法很有意思，但关于它如何发生的详情却从不为人所知，而且可能还有其他更有趣的发生方式。假设你的大脑被移植于另一个新身体之中，它们彼此会配合默契吗？你是不是会认为，这等于换了个身体？当然，这里有无数技术问题，但是，就我们的目的而言，我们可以不讨论这些问题。

假如你的大脑被置入另一人的身体，你会同新伙伴唇齿相依，是不是？但问题是，你是一个大脑吗？你试试下面两个句子，看看对你来说哪句更正确：

我有一个大脑。

我是一个大脑。

我们平时在谈到一个人特别聪明时，常将他比作大脑，但这只是一个比喻。我们的意思是说，他们的脑子很好。你有一个好使的脑子，但具有这个脑子的你究竟是谁或是什么？还有，如果你有一个大脑，你能拿它换另一个大脑吗？如果你在身体调包时总是和你的大脑在一起，那有谁能够在大脑调包时将你和你的大脑分开来呢？不可能吗？难说。过一会我们就会看到。

总之，如果你刚从火星上回来，你已将你以前的大脑丢在那里了，对不对？

现在假定我们同意这样的假设，即你有一个大脑。你以前是否曾自问过，你怎么知道你有一个大脑呢？你从来没看到过它，对不对？哪怕面对着镜子，你也看不到它；你也摸不到它。但是，尽管如此，你还是知道你有一个大脑。你知道这一点是因为你知道你是一个人，而凡是人都有大脑。凡是人也都有肝，但奇怪的是，你知道自己有大脑和有肝的途径竟是一样的。你相信从书中获得的知识。人们以前一直不知道肝的功能，是科学发现了这个答案。人们也未必都知道大脑的功能。曾有过这样的说法，亚里士多德认为大脑的功能是冷却血液，——当然在它的运动过程中，它确实十分有效地冷却了你的血液。假设我们的肝长在脑袋里，而大脑却被移植到胸腔内。当我们举目观望，侧耳倾听时，你会不会认为我们用肝思维的说法也很有道理？你的思维看上去发生在眼后耳间——但这是因为大脑长在那个位置上，还是因为你将你自己放在你的目力所在的那个地方？试图设想如何能用大脑——白嫩的花椰菜般的东西去思维，与设想如何能用肝脏——红红嫩嫩的东西去思维，事实上不正好是一样使人触目惊心吗？

你是什么不只是个有生命的身体(或有生命的大脑)，而且也是灵魂或精神这个概念，尽管它具有古老的传统，但在许多人看来是不科学的。他们会这样说："灵魂在科学中没有立足之地，永远也不能为科学的世界观所接纳。"科学使我们懂得，从来不曾有过灵魂这样的玩艺儿。我们不再相信妖魔精灵。多亏了科学，那种认为身体内有灵魂——"机器中的幽灵"——的值得怀疑的观点很快就自行将幽灵拒之门外。关于你同你纯物质的身体有区别的观点有许多种说法，有的说法不见得那么不堪一击。

我们会看到，有些说法吐蕾于科学的花园中。

我们的世界充满万物，万物并不神秘，但也不像物理学的砖块机械垒叠得那么简单。你相信声音吗？你对理发有何感想？有这类事物吗？那是些什么东西？用物理学家的语言说，什么是洞——不是指一个异常的黑洞，而只是指例如一块奶酪里的一个洞。这是物质的东西吗？交响乐是什么？“星条旗”存在于时空中的哪个地方？它不就是国会图书馆里某一张纸上的墨迹吗？烧了这张纸，美国国歌依然存在。拉丁语还存在着，尽管它不再是有生命的语言了。法国洞穴人的语言早已不存在了。桥牌的历史还不到一百年。它是什么？它不是动物，不是蔬菜，也不是矿石。

这些玩艺儿都不是有体积的物质，也不是化合物，但也不是纯抽象的东西——像数学中的π，那是个不能存在于时空中的不可改变的数，这些玩艺儿有出生地和历史。它们自己会有变化，也会因外来影响而改变。它们能够像物种、疾病和流行病那样传播。我们绝不能认为，科学告诉我们的仅仅是，世上被任何人所认真看待的万物都是由在时间和空间中运动的分子集合所构成的。有些人认为你不过是一个特殊的物质生命有机体——一堆移动的原子——这不过是常识（或只是正常的科学思维）而已，但实际上，这个观点表明，人们缺乏的是科学的想象力，而不是冷静的思考。我们用不着非得先去相信幽灵，然后才相信有超越任何特殊生物体特性的自我。

你毕竟还是萨拉的母亲。但萨拉的母亲是你吗？她是死在火星上了呢，还是又重返地球了？在你看来她已归来了——当然，在她跨入运输舱之前，在她看来她将重返地球。她这样想对吗？也许对。但你对使用这种新型马克4型长途传真运输机的结果有何高论呢？多亏了非侵袭性计算机辅助测试扫描技术产生的

奇迹，使运输机在没有毁灭原本的情况之下，获得了蓝本。在火星上，萨拉的母亲还能够选择是去还是留。她能拨亮电钮，走进运输舱，那是为了萨拉的缘故，也是为了能回到地球用一位能言善辩的女发言人的言辞向人们讲述整个发生在火星上的悲剧；她也能够走出这运输舱，继续留在火星上。一个人是不是真的能同时处在两个地方？不管怎么说，这样的情况不会长久，因为分身两处很快就会积累起不同的记忆和不同的经验，会成为截然不同的两个人。

私人生活

是什么使你成为你，你的界限是什么？答案的一部分很清楚——你是意识的中心。但意识究竟为何物？意识是心灵最明显也是最神秘的特征。一方面，我们每一个人对自己的经验、知觉、痛楚、观点和思想了如指掌，不容置疑。另一方面，意识到底是什么东西？生活在物质世界内的物质生命怎么会制造出这样一种现象？科学揭示了许多一开始很神秘的现象的本质，像磁力、光合作用、消化以及生殖。但是意识与它们不尽相同。原因只有一个，即磁力、光合作用和消化的具体情况在原则上都能被配备合适仪器的观察者所发现，但是，熟悉有关意识的任何具体情况似乎只有一位有天赋的或有特权的观察者。他认识意识的途径完全不同于并且胜过任何其他人的途径，无论他们可能配备什么样的仪器。由于这些和另一些原因，至今还没有一个较好的关于意识的理论。甚至关于这个理论是什么样子，也还没有一致的看法。因此有些人干脆就否认“意识”这个词有真实的所指。

长期以来，我们生活中这样一个尽人皆知的特征一直没能为人们所识透。单单这个事实就能说明我们关于意识的观念是错的。现在需要的不是更多的证明，也不是更多的实验和临床的数据，而是要对种种假设进行反思，这些假设导致我们认为存在着单一、熟知的意识现象，它能与意识一词的日常含义所允准的一切描述相符。考虑一下在注意意识时所必然要提出的一些使人感到迷惑的问题。其他动物有意识吗？如果有，它们的意识与人的意识是否一样？计算机和机器人有意识吗？人是否有无意识的思维？是不是还有无意识的疼痛、感觉或知觉？婴儿在出生之前和出生之时有意识吗？我们在梦中是否有意识？一个人能在一个大脑里具备数个意识主体或自我或动因吗？要对以上这些问题作出较满意的回答，肯定必须对意识的几个成问题的取代者的行为能力和内部状况作经验的研究，而对每项这样的经验研究结果，我们都要提出这样一个问题：它对意识的问题有什么意义？为什么？这些并非是直接的经验问题，而是观念问题，观念问题也许可以通过思维实验而获得解释。

我们对意识的一般看法主要有两大类，一类可以用“从内部”来概括，另一类则可用“从外部”来概括。从内部看我们自己的意识似乎很明显，无处不在：我们了解周围许多事物，甚至也知道我们完全没有感觉和意识到的体内情况，可我们对自己的意识的了解却是最清楚不过了。我所意识到的那些事情，以及我意识它们的种种方式，这两个方面决定了我是什么？我是以一种人所没有的方式来认识我是什么的。从内部看，意识是个非此即彼的现象——好比心中的一盏灯，或是亮着或是熄灭。我们承认，我们有时候昏昏欲睡，或心不在焉，或麻木不仁，有时我们甚至会处在反常的强意识之中。可是当我们有意识时，有意识这个事实就不允许有程度之差。有就是有。没有

就是没有。有一种看法认为，意识似乎是将宇宙分成两个截然不同的世界的特征，一个世界有意识，而另一个却没有。有意识的就是主体，对于主体事物会有不同表现，同时主体也不是一般的物质。一块砖头，一架袖珍计算器，或者一只苹果，这些都不是主体。它们有内在结构，但不是那种结构——不是内在生命，也没有观点，我当然是一个主体(我“从内部”知道我是主体)，你同样也是主体(因为你已告诉过我，你也有同感，这非常有说服力)，也许狗或海豚也是主体（要是它们自己能告诉我们就好了！)，甚至蜘蛛也是主体。

他人之心灵

当我们想到他人时(包括同类或其他生物)，我们是从外部认识他们的，因此，我们觉得他们各种各样可观察到的特征都与他们的意识问题有关。生物在它们的感觉范围之内对事件作出恰当的反应。它们识别事物，避免引起痛苦的环境，它们学习、计划、解决问题。它们表现出智力。但是，用这种方式讨论问题，也许会有偏见。比方说，一开始就谈论生物的“感觉”或“痛苦”的情境，就等于说我们已经解决了意识这个问题——因为，请注意，要是我们用这种方式来谈论机器人，这种有疑问的选词意图就会一目了然(许多人对此有异议)。生物同机器人有何区别，区别是真的还是臆想出来的?它们的机体和生物结构同我们相似，而我们是典型的有意识的生物。当然，这种相似性有不同的差异，至于我们直觉地计算这种相似性的差异究竟有多大，这往往是不可靠的。海豚的属鱼性使它们不可能和我们人类有同样的意识，毫无疑问也不应该是一样的。即使大

猩猩同海参一样迟钝，可由于它们同人类有表面的相似之处，也会使它们被列入高级智能生物的行列。如果家蝇的体积同人类一样庞大，或者它们是热血动物，在我们折断它们的翅膀时，我们就一定会以为它们感到了疼痛（是我们感觉的那种疼痛，这里讲究的是种类）。是什么使我们认为有的考虑有价值，而有的则没有呢？

对此有很明显的答案：各种"外部"特征多少都是每个意识主体从内部认识到有某种东西存在的可靠迹象或症状。但怎样证明这个答案呢？这就是著名的"他人心灵之谜"。拿一个人来说，他能够直接观察到自己的内在生命与外在可观察行为的一致性。但是，如果我们每一个人都想严格地跨越唯我论，那就必须做一件表面看来不可能的事：证明他人的内在生命与外在行为的一致性。严格说来，光是听他们说他们自己有这种一致性是不行的，因为这只能提供给我们更多的外在行为与外在行为之间的一致性：知觉和智力行为的可观察能力一般总是与语言能力尤其是"内省"陈述能力一同发展的。如果一个设计巧妙的机器人（看上去）能够向我们叙述它的内在生命（在适当的环境里说适当的话），我们可否将它列入高级智能生物？我们可以这么做，但我们怎么能保证没有受骗？这里的问题是：内心的那盏特殊的灯是真的亮着，抑或根本没有什么灯，只是一片漆黑？这个问题似乎无法回答。所以，我们也许已经走错了一步。

我在这几段话里关于"我们"和"我们的"的用法，以及你们不加怀疑地将它们接受下来的态度，都说明我们并没有认真思考他人之心灵这个问题——至少对我们自己来说，对我们与之正常交往的人类来说是这样的。人们很容易认为，如果还有关于臆想的机器人（或是某个成问题的生物）的重大问题没有得到解决，那必定是可以通过直接的观察而得到解决的。某些理论

家认为，一旦我们对大脑的结构和它们控制行为的作用形成了更好的理论，那么，我们就可以运用这些理论来区别意识实体和无意识实体。这就是说，通过某种方法，就可以把那些我们个人"从内部"获得的事实还原为可从外部公开认识的事实。从外部公开认识的事实达到一定数量，就能够解决某些生物是不是有意识的难题。比方说，让我们来考察一下神经生理学家E·R·约翰的说法，他试图用客观的语言来给意识下定义。他说：意识

> ……是这样一个过程，在这个过程里，有关感觉和知觉的各种各样个体样式的信息结合起来构成了对该系统及其环境的状态的一种统一而多元的表现，它同关于记忆和机体需要的信息相结合，从而产生了行为的感情反应和程序，以调整机体适应它所处的环境。

要想确定这个假设的内在过程发生在某个特定的机体内，这在关于神经信息处理的新科学领域里，可能是困难的，也是个依赖于经验的使命。我们假定，在某个特定生物体内，这个过程成功地完成了，那么，基于这一点，可以说这个生物是有意识的。如果我们准确地理解了这个观点，也就没有什么值得再疑神疑鬼的了。在这里，如果持保留的态度，那就好比在有人领着你仔细地观看了一台汽车引擎的工作状况以后，你还问道："但是，这真的是一台内燃机吗？我们没有受骗上当吧？"

任何关于意识现象的正确的科学解释，最终都必须采取这样一个教条式的步骤，即要求这个现象能够被客观地观察到，但是，这里面还有一个问题，一旦采取这样一个步骤，那个神秘的现象是否就被抛在脑后了。在将这种怀疑论的预感作为浪漫主义者的狂想而摈弃之前，我们不妨先考察一下当代心灵研究

史中发生的一场惊心动魄的革命，这场革命带来了不可平息的影响。

弗洛伊德的拐杖

对于约翰·洛克和继后的许多思想家们来说，意识，特别是自我意识在心灵里起着至关重要的作用。心灵在它的一切活动和过程中对它自己来说是一清二楚的，从内心的眼睛看来，一切都清清楚楚。为了明察内心的活动，你只要"看"就行了——也就是你"内省"——你由此而发现的界限也正是心灵的界限。人们没有接受潜意识思维或知觉这个概念，即使一旦接受了它，它也将会因为被认为是不连贯、相互矛盾的废话而被摈弃。在洛克看来，有一个十分棘手的困难，即怎样将"在意识面前"不连续存在的记忆描述成在人心中是连续存在的。这个观点影响巨大，直至后来当弗洛伊德最初假设存在着潜意识心理过程时，普遍遭到了直截了当的否定和诘难。这不仅是对常识的触犯，也是自相矛盾的，因为它认定存在着潜意识的信仰和欲望、潜意识的仇恨、潜意识的自卫和报复的谋划，但是，弗洛伊德最后却赢得了反对者的拥护。一旦出现下述情况，理论家们就发现这种"概念的不可能性"是可以想象的，因为它允许人们去解释许多不然就无法说明的心理病理学的模式。

这种新的思想方式依靠着一根拐杖：我们至少能够抓住一个洛克主义信条中苍白无力的说法，我们可以这样想，这些"潜意识的"思想欲望和谋划在精神上属于他人的自己。正和我能够将自己的谋划隐藏起来不让你知道那样，我的本我同样也能不让我的自我知道。通过将一个主体分割成数个主体，我们就

能够维护这样一个公理，即每一个心理状态必定是某人的有意识的心理状态，通过假设某些心理状态另有一些内在的拥有者，我们就能解释为什么它们不为其假定的拥有者所察知。这个步骤在术语的雾霭里完全模糊不清了，这反而因祸得福，比方说，什么是超我这样一个令人费神的问题，可以被搁在一旁，置之不理。

弗洛伊德对可思想事物界限的扩充，彻底革新了临床心理学，它也为更晚近的“认知”实验心理学的发展铺平了道路。我们现在毫不犹豫地接受了一大堆说法，我们坚信，复杂的假设检验、记忆搜寻和推论，一句话，即信息处理，就发生在我们体内，尽管对于这一点通过内省是无从知道的。不是弗洛伊德所发现的那种被压抑的潜意识活动，即被逐出意识“视野”的活动，而仅仅是心理的活动深埋在意识下面，或是完全为意识所觉察不到。弗洛伊德声称他的理论和临床观察使他有权力驳斥病人对自己内心思想的自信的否认。同样，认知心理学家也可以列举种种实验证明、模式和理论，以表明人们有着极复杂的推理过程，然而对之却根本无法作内省的说明。因此，心灵不仅可被外人所知，而且，有些心理活动在外人看来比在它们真正的“主人”看来更清楚。

然而，在一些新的理论中，这根拐杖被扔掉了。在这些新的理论中，充满着精心设计的、富于幻想的小矮人的比喻，子系统被比作大脑中的小矮人，他们传递信息，寻求帮助，服从命令，自告奋勇。尽管如此，真正的子系统被认为是机体机器中没有问题的非意识部分，就像肾和膝盖骨那样没有主见，也没有内在生命（当然，“无心灵”的却有“智能”的机器人的出现，对进一步消除洛克式的观点起了很大的作用）。

但是，现在洛克的极端主义已被驳得体无完肤。因为，如

果潜意识心理这个观念被认为是不可理解的，那么，我们也就无法把握住有意识心理这个观念。要是完全潜意识的、无主体的信息处理过程在原则上能够做到有意识心灵所能做到的一切，那么，意识还有什么用呢？要是认知心理学理论对我们成立，那么，它们对鸡尾酒和机器人也能成立。这样的话，这些理论就无法辨认我们是谁了。单纯的无主体信息处理过程（即最近在人体中发现的那种过程）的总量能成为与之成鲜明对照的特殊特征吗？因为两者之间的这种差别至今尚未消失。心理学家卡尔·拉士利曾用挑衅的口吻说道："心灵中没有什么活动是有意识的。"他的话意在让人们注意，我们对自己思想时所发生的信息处理过程是无从知晓的。他举了个例子：要是你问什么叫扬抑抑格六韵步，知道的人就会马上告诉你。比方说：我究竟是怎样理解扬抑抑格六韵步这个情况的？我们怎样作扬抑抑格六韵步？产生这样一个思想时，我们的大脑有什么活动？这些问题我们难以回答。拉士利的说法乍看上去是在说，意识这个现象是由心理学来研究的，实际上，他的真正意思正好相反。他的话使我们正确地注意到所有潜意识信息处理过程与可以直接认识的意识思想本身之间存在着差异；毫无疑问，如果没有这种潜意识信息处理过程，就没有意识经验。但这种意识思想由什么或由谁来认识？如果由大脑的某些子系统来认识，那就没有将意识思维与潜意识活动和事件区分开来，而潜意识活动和事件也可以被大脑的各种子系统所认识的。如果假设大脑中有一个子系统其结构非常特殊，它与该系统的其他方面相互关联而构成了世界中的另一个自己，另一个"有某种特征的东西"，这是很不好理解的。

但是很奇怪，这是一个陈腐的问题，既然认知科学已将人心分析成种种功能构成，他人之心灵的问题就重又作为一个严

肃的问题出现了。这在分割大脑例子中表现得最为突出。我们现在可以肯定地说，经过大脑胼胝体切割手术的人会有两个相对独立的心灵，一个同主要的脑半球相联，而另一个则同次要的脑半球相接。这不会有什么问题，因为我们已习惯地认为，一个人的心灵是由相互沟通的亚心灵构成的。在这里，沟通的渠道完全被割断了，这样，每一部分独立的特性就鲜明地显示出来了。但是，这里还有另外一个问题：两个亚心灵是不是“有它们各自内在的生命”？一种看法认为，次要的脑半球不会有意识，因为就我们所知，这部分脑半球同别的许多潜意识的认知子系统一样，能够处理许多信息并能以智能控制某些行为。但是，我们不禁要问，为什么说在主要脑半球里就会有意识呢？此外，为什么在一个常人完整的、未受损伤的脑系统里就有意识了呢？我们以前曾认为这个问题不值一论，但是，现在这种状况迫使我们重新正视这个问题。另一方面，如果我们坚持说次要的那部分脑半球（说的更恰当一点，是新发现的那个长着次要脑半球的人）有“具备内在生命”的意识，那么，对于当今理论中假设的所有其他信息处理子系统，我们又该作何评价呢？我们是不是可以不惜以增加许多头脑，其实是增加大量的经验主体为代价，来重新拾起弗洛伊德的那根拐杖？

现在，让我们看一下心理语言学家詹姆斯·拉克纳和梅里尔·加勒特的令人惊叹的发现。这个发现可以称为句子理解的潜意识声道。在一个双声道的听力测验中，被试通过耳机可以听到两个不同的声道，但只要求他们倾听一个声道。一般他们都能准确地解释或汇报他们注意听的那个声道中的句子，但对于另一个同时传送但没被注意听的声道中的句子他们一般都说不出什么来。因而，如果在那个没被注意听的声道里传送了一句话，被试一般都能说他们听到了一个声音，甚至能说出那声音是

男性的或是女性的。也许，他们甚至还能确定那个人说的是不是他们的母语，但他们说不出他说了些什么。在拉克纳和加勒特的这个实验里，被试在侧耳倾听的那个声道中听到了一些意思很难琢磨的句子，比方说，像"他吹灭灯笼报警"。在没有注意听的那个声道中，一组被试同时也听到了一个句子，这个句子实际上解释了另一个被注意听的声道中的那个句子(比如，"他熄灭了灯笼")，而另一组被试却听到了一个中性或是没有关联的句子。前一组被试说不出来在那个没有注意听的声道里讲了些什么，但他们比对照组的被试更容易接受那种对于意义含混的句子的提示性解释。毫无疑问，从没有注意听的声道中传出的信号对于注意听的信号的解释有一定影响，要解释这一点只有作这样的假设，即没有被注意听的信号一直是在语义层次上被处理的——换句话说没有注意听的信号被理解了——但这很明显是潜意识的句子理解！或者，是不是我们可以这样说，这是证明被试体内至少存在着两个不同的、仅有局部交流的意识？当我们询问被试对理解没有注意听的信号有什么体会时，他们会诚恳地说，没有什么体会。实际上他们并没有注意那个句子。但是，就像在大脑割裂的病人中经常看到的那样，我们提问的对象其实也许是另外一个人，这位被试有意识地理解了那个句子，并将它的某种含义暗示给那个回答我们提问的被试。

我们该认为哪个解释有道理呢？为什么？我们似乎又回到了那个不可回答的问题，这个问题建议我们应该从不同的角度来考察这种情况。如果要形成一种能对众多的难题作出公平评价的意识理论，那就需要在我们的思维习惯内发动一场革命。破除坏习惯并不那么容易。本书收集的奇思冥想和思维实验都是用来帮助破除坏习惯的游戏和练习。

博尔赫斯和我

豪尔赫·路易斯·博尔赫斯

事情发生在另外那个叫作博尔赫斯的人身上。我漫步在布宜诺斯艾利斯的街头，不时止步观望，也许现在正在不由自主地打量拱形的门和大门上的铁花格；我是从邮件中得知博尔赫斯的，并且从一串教授的名字或一本人名录中看见了他的名字。我喜欢古代计时沙漏、地图、十八世纪的印刷术、品尝咖啡和欣赏史蒂文森的散文；他也有这些爱好，但不是演员的那种假装的爱好。要是说我俩是一对冤家，这也许有点夸张；我活着，一直活下去，这样博尔赫斯就可以舞文弄墨，而他的文墨又为我提供了存在依据。我用不着否认他写了几页有用的文章，但这几页文章拯救不了我，也许是因为好的东西不属于任何人，甚至也不属于他，而是属于语言和传统。除此以外，最终我肯定会消亡，只有我生命的某一刻会残存在他的身上。我一点一点地把自己的一切都移交给了他，尽管我很清楚他有说假话、说大话这样的坏习惯。斯宾诺莎知道万事万物都希望长生不死；石头永远希冀自己是一块石头，而一只猛虎也同样期望自己永远是

猛虎。我将在博尔赫斯身上，而不是在我自己身上活下去（如果我真是某个人的话），但我总是在别人的书中或是吉他弹奏声中，而不是在他写的书中发现我自己。几年以前，我试图摆脱他，因此，我就动笔写作，从郊外的神话写到有着时间和永恒的游戏，但那些游戏现在属于博尔赫斯了，我不得不再去构思别的故事。就这样，我的一生像一场旅行，来去空空，我失去了一切，而一切将被遗忘，或属于博尔赫斯。

我不知道是我还是博尔赫斯写了这些话。

D·C·丹尼特

D·R·霍夫施塔特

反　思

豪尔赫·路易斯·博尔赫斯，这位伟大的阿根廷作家享有国际声誉，而这声誉却带来了奇特的后果。博尔赫斯在他自己看来宛如两个人，一个是众所周知的博尔赫斯，而另一个是鲜为人知的博尔赫斯。他的名望又扩大了这种后果，但是，正如他所知道的那样，我们都有这种感觉。你在一个名单上看到了自己的名字，看到一张你自己的传真照片，你听到人家在议论某人并且突然意识到被议论的那个人就是你。你的心灵必定要从第三人称的角度——“他”或“她”——跳到第一人称的角度——“我”。喜剧演员早就知道了如何夸张这种跳跃。这也就是典型的“先是一怔，后来才恍然大悟”。比方说，鲍勃·霍普[①]

① 鲍勃·霍普，美国好莱坞著名喜剧演员。——译者

在读晨报时看到鲍勃·霍普被警察通缉，心不在焉地评论了几句，然后突然蹦了起来："那是在说我！"

罗伯特·彭斯曾经说过，像别人看到我们那样看到我们自己，是一种天赋，这也许是对的；尽管如此，这决不是我们能够或者应该始终追求的境界。事实上，近来有几位哲学家已经发表了高明的论断，说明在认识我们自己这个问题上存在着两种完全不同的和不可简化的方法。这些论点的技术性很强，但讨论的问题非常有意思，并且能以生动的例子加以阐述。

彼得正在商店里排队付款，他发现柜台上方有一架闭路电视——这是商店为了防止小偷偷窃所采取的措施之一。他在观看荧光屏上拥挤不堪的人群时，发现荧光屏上左边那个身着外套提一个大纸袋的人的口袋被他后面的那个人摸了。他抬手吃惊地捂住了嘴，他看到那位被窃者的手也同样伸向他的嘴。彼得突然意识到他就是那个被窃者！这个戏剧性的变化是一个发现；彼得慢慢懂得了他几分钟以前不明白的事情，当然，这很重要。要是没有那种能激起他采取防卫行动的思想，他就几乎不能采取任何行动。但在这变化以前，他当然也不是漠然无知的，他正在思考那个"身着外套的人"，看到那个人被窃了，因为那个穿外套的人就是他自己；他刚才正在思考**有关他自己的事**。但是，他并没有把自己**当作自己**来思考，也就是说，他没有正确地"思考自己"。

另外再举一个例子。设想一下某人在读一本书，书上有一段话，其中第一句话用一个由几十个词构成的描述性名词短语描写了一个最初还不知性别的无名人士在干一件平凡的事。他读了这一短语之后，在他或她的想象中马上形成了一个简单而模糊的心理意象，即某一个人在干某件平平淡淡的事情。再往下读几句，随着描写的细节不断补充，这位读者对整个故事情

节的心理意象逐渐集中在某个比较明显的焦点上。接着，在某一刻，由于描写变得相当具体而使这位读者心中一颤，骤然不安起来，因为他或她自己就是书中所描写的那个人！“我真傻，居然没早看出我在读关于我自己的书！”这位读者陷入了沉思，感到有点局促不安，但同时也相当兴奋。你也许能想象这样的事发生，但为了帮助你更好地理解，可以假定上面谈到的那本书就是《心我论》。难道你现在对整个故事情节的心理意象还没有达到一个较明显的焦点吗？它是不是突然教你心中发颤？你想象读者在读书中的哪一页？哪一段？读者心中会闪过什么思想？如果读者是个真人，那他或她现在正在干什么？

要描写一种能对自己作如此自我表述的事物是不容易的。假定一架计算机能通过程序来控制以无线电链路与其相连的机器人的运动和行为（在加利福尼亚的国际斯坦福研究所的那个叫做“沙基”的机器人就是这样控制的）。计算机贮存着机器人及其环境的表象，当机器人活动时，这种表象也随之变化。这样，计算机程序就能借助关于该机器人的“身体”和所处环境的最新信息来控制着这个机器人的活动。现在假定，计算机将机器人表示为位于一间空屋的中央，你呢，则被要求将该计算机的内在表象“翻译成英文”。那么是“它（或者他或沙基）站在空屋的中央”还是“我站在空屋的中央”？这个问题将改头换面重新在本书讨论视心灵为程序的观点部分出现。

2 无头有感

D·E·哈丁

我一生中最美好的一天——比如说我的再生之日——是我发现自己没有头的那一天。这并不是卖弄文学技能、不惜代价哗众取宠的俏皮话。我是在正儿八经地说这句话：我没有脑袋。

我是18年前，在我33岁时发现这一点的。尽管这个想法是心血来潮，但也并非无缘无故。好几个月以来，我一直在思考这样一个问题。我是什么？当然，我当时正巧在喜马拉雅山漫步，与这问题也许并没有联系；尽管人们认为在这样一个地方，极易出现非凡的心理状态。那天，天高云淡，我站在山脊上远眺，面前是云雾弥漫的蓝色山谷，直插苍穹的世界最高峰，干城章嘉峰和珠穆朗玛峰在雪峰中隐约可见。不管怎么说，这是一个令人产生伟大思想的地方。

实际上发生的事情是极其简单和渺小的：我停止了思考。我只觉得有一种奇怪的静止感，一种奇异的、又警觉又无力的或麻木的感觉。理智、想象和其他一切心理活动都消失了。我一

时说不出话来。过去和未来都逝去了。我忘却了我是谁、我是什么，我的姓名、人性和动物本能，凡是我的一切都被忘却了。好像我是在那一瞬间出生的，焕然一新，头脑空空，忘记了过去的一切。只有现在存在着，只有当下的这一刻和在这一刻中清晰可见的东西存在着。你只要睁眼看就是了。我只见，卡其布裤子下面有一双棕色的鞋，两只卡其布袖子下各有一只粉红色的手，卡其布衬衫上端——其实什么也没有！当然没有脑袋。

我顿悟到这所谓的什么也没有的地方，本该有一个脑袋的那个窟窿，并不是一般的空缺，并非单纯的虚无。相反，那里包罗万象。这是一个无比充实的空缺，这是能容纳一切的虚无——草地、树林、隐约的远山，以及在这一切之上如同蓝天上一排有棱角的云朵一样的雪峰。我失去了一个脑袋，但获得了一个世界。

这的确惊心动魄。我真觉得呼吸停止了。整个身心都沉浸在这造物之中。这片壮丽的景色，在晴空中闪耀，不凭什么依托，神秘地悬在半空中（这是真正的奇迹、神奇和快乐），它完全不受"我"的支配，不受任何观察者的影响。它的完全存在也就是我肉体和灵魂的彻底消失。它比空气还轻盈，比玻璃还明澈，完全脱离了我自己，我哪里也不存在。

尽管我所看到的景色具有神秘甚至离奇的色彩，但这不是梦，也不是不可思议的启示。恰恰相反，这好比是从普通生活的睡梦中突然惊醒，是一场美梦的结束。这是从模模糊糊思想状态中一下子解脱出来的自明的现实。这归根结底是对显而易见的事物的启示。这是混沌的一生中唯一清晰的一刻。这时不再对过去由于（无论如何是从我孩提时代以来）太忙或太聪明而没有发现的东西视而不见了。这是对过去一直呈现在我面前的事物投以明白简单、毫无批判目光的时刻——我成了个毫无个

性的人。总而言之，一切都非常简单、明白、直截了当，不必争论，不必思维，也不必说话。没有疑问，也没有经验以外的东西，只有宁静和静谧的愉悦，只有如释重负的轻松感觉。

* * *

随着我对喜马拉雅山上的发现表现出的最初的惊讶逐渐消逝，我便开始向自己描述当时的情景。下面就是我的描述。

我以前隐隐约约地总觉得我自己是居住在叫作身体的房子内，透过两个圆圆的窗户窥视着世界。现在我发现绝非如此。当我凝视远处时，可根据什么来告诉我，我有多少只眼睛？——一双、三只、几百只，还是一只也没有？实际上，我的正面只有一个窗口，而这个窗口是开着的，没有窗框，窗内并无人向外张望。长着一双眼和有一张脸的总是另外一个人，而不是这个人。

因此，人可以分成截然不同的两大类型。第一类人不计其数，肩上扛着一个脑袋（我这里说的“脑袋”指的是上面有几个孔、有毛发的8英寸的球体），而第二类，我只见到一个，两肩上显然没扛着这样的玩艺儿。我居然会一直都没有注意到这么大的区别！我是持久愚蠢的牺牲品，也是终生幻觉的牺牲品（我这里说“幻觉”指的是我的词典解释的意义：即对某种实际上不存在的物体的表面感知），我常觉得自己同别人并无什么差别，当然不以为自己是砍了脑袋还活着的两足动物。我对一直存在着的那一件东西熟视无睹，若没有它我真的会变成盲人——它就是令人惊叹的脑袋的代替物，它清澈无比，它是明亮而又纯洁的虚空，然而它就是——确切地说不是包含——一切。因为，无论我多留神，我在这里竟不能发现显现群山、太阳和天空的屏幕，没有发现反映这景象的明镜，没有发现可以透视这

景象的透镜或者视孔——更不用说有接受这景象的心灵，或是与景象本身不同的观景者(尽管多么不真实)。没有一点干扰的东西，哪怕是那个捉不住摸不着的被称作“距离”的干扰物。辽阔的蓝天、镶着红边的雪和碧绿的草地——既然无所谓远与近，那么所有这些怎么可能是遥远的？在这种情况下，这个无头的虚空既不能描述，也无着落；它不是圆的，不是大的，也不是小的，甚至不在(与彼地相异的)此地(即使真有一颗向外伸的脑袋，从它向珠穆朗玛峰伸展的测量杆，如果从一端开始读数的话——我只能这样读——将降为零，降为乌有)。事实上，这些五彩缤纷的物体以最简单的方式呈现出来，没有任何诸如远近、那个这个、我的你的，以及被我看见或是只是存在之类的复杂概念。一切二元的东西，即主观和客观的双重性，全然不存在了；这种双重性已经不能加进不容纳这种东西的状态里去了。

这就是我即景而发的感想。但是，想用这样或那样的文字来描述这样第一手的、直接的经验，都会因为将十分简单的事情复杂化而得出错误的结论；事实上这种事后的思考越是拖得久，越是与真实的情形大相径庭。这些文字描述至多只能使人回忆起当时的情形(却不再那样栩栩如生)，或者使当时情形重现；但它们无法重现其本质，或确保其重现，这种情形就与令人垂涎的菜单不会有佳肴的美味，最好的幽默论著不能使人体察一个笑话一样。另一方面，人们不可能长期停止思考，有时候人们往往会将人生中某些清晰的时刻与混乱的背景联系起来。这样的话，也会间接地导致清晰时刻的重现。

不管怎样，仍然存在着若干个不愿被敷衍的常识性异议，也即存在着若干个要求得到合乎逻辑的尽管是不全面的答案的问题。因此，这对于“确证”一个人的视觉是否正确，即使对他自己来说也是必要的，他的朋友们或许还需要重新确证。从某种

意义上说，这种归化的做法是荒谬的，因为对于确定明了、无可争议的经验添加或删减任何论证都是不可能的，就像中音C便是中音C，草莓酱便是草莓酱一样。而从另一个意义看，又非得如此，要不然的话，人的生活就会被肢解成互相排斥、无法疏通思想的两大部分。

*　　　*　　　*

我第一个异议是：我的头也许掉了，可鼻子还安然无恙。无论我走到哪里，它都在我面前。对此我的回答是：如果我的右边悬着一朵模糊而又完全透明的粉红色的云，左边悬着同样的一朵云，它们都是鼻子的话，那么，我就有两个鼻子，而不是一个；而我清楚地看到在你脸中央隆起的那块完全透明的肉赘并不是一个鼻子：除非是一个完全不诚实或是全然糊涂的人，才会用同一个名称去指代两样截然不同的东西。我倒情愿依我的字典和普通词义办事，那我不得不这样说：尽管几乎所有的人各有一个鼻子，我却没有。

同样，如果某个误入歧途的怀疑主义者，迫不及待地想申明自己的观点，对准这两个粉红的云朵中间打来，那么，我肯定结果一定是不愉快的，就像要是我长了一个最坚固、最耐打的鼻子遭了打也会不愉快一样。还有，该如何解释从来没有完全脱离这块中央地区的种种感觉，比如微妙的紧张、运动、压力、发痒、疼痛、温暖、颤动等等？首先如何解释当我伸手抚摸这里时所产生的触感？是不是说，这些发现也可以作为大量证明此时此地我脑袋的存在的证据？

绝对不能。很明显，这里罗列了许多不容怀疑和不容忽视的感觉，但它们不能相当于一个脑袋，或相当于类似脑袋的东西。要想通过这些感觉证明有一个脑袋，唯一的办法就是拿

出各种各样的这里显然缺乏的因素来——尤其是各种立体的、有色的因素。要是一个脑袋没有像其他脑袋那样有眼睛、耳朵、嘴巴、头发以及凡是脑袋所必须具备的一切，光是有各种各样的感觉，这算是个什么样的脑袋呢？事实显而易见，这块地域不能有丝毫的阻碍，丝毫的含混和杂色，要不然，我的世界就会烟雾笼罩。

无论怎么样，在我伸手去寻找我失去的脑袋时，非但没有找到它；反而把手也丢了，它在我自身存在的深渊中被吞噬了。很明显，这块我以为长着脑袋的地方，这个开着大口子的洞穴，这块我一切行为发源的空旷基地，这块神奇莫测的地方，实际上更像熊熊燃烧的灯塔，把旁的一切都燃烧了，因而，这照亮世界的光芒将永不暗淡。至于那些潜在的痛痒之类的感觉，它们也如群山、乌云以及天空一样，不能遮挡这明亮的光辉。相反，它们都沐浴在这片光明之中，通过它们可以看到灯塔的灿烂光芒。现在对世界的体验，不管从什么意义上来理解都发生在虚无的、不在的脑袋之中。因为此时此地，我的世界与我的脑袋是不共戴天的，是决不相容的。人的肩膀上不可能同时容许世界和脑袋并存，值得庆幸的是，我的脑袋及其一切构造必须离去。这一点无需争论，无需哲学上的敏锐，也不必说服人，只需要纯粹的直觉：是看一看是谁在这里，而不是想一想是谁在这里。如果说我看不见我是什么(尤其是我不是什么)，那是因为我太热衷于想象。太“讲究精神”，太老于世故了，因此无法完全接受我现在所发现的处境。我所需要的是大智若愚。只有天真的眼睛和虚空的脑袋才能发现它们自己的彻底的空虚。

* * *

要想使一位声称“我在这里有一个脑袋”的怀疑主义者改变

态度，恐怕只有一种办法，那就是邀请他到这里来，让他自己瞧一瞧；可他必须诚实，如实描述所观察到的事情，如此而已。

在屋子的另一端，他看到我是一个长着脑袋的人的全身形象。但是，当他朝我这边走来时，他发现我成了半个人，然后只有一个脑袋，然后只看见模糊的面颊、眼睛或鼻子，然后只见一片模糊，最后（在接触的一瞬间）什么也没有了。要是他碰巧带着必要的科学仪器，他说那一片模糊又变成了皮肤组织，然后变成细胞群，然后是一个单个细胞，然后是细胞核，然后是大分子……如此下去，最后他来到一个什么也看不见的地方，来到没有任何固体或物质客体的空间。在以上这两种情况中，这位到此进行探索的人看见的事情与我的所见相同——都是虚空。如果他发现我不在此处，就会环顾四周（与我一同朝外看，而不是盯着我看），结果也会看到我所见的东西，即这个虚空充满了凡能想象的一切。他还会发现这个中心点扩展为无限的量，这个无变成了一切，这个此处变成了到处。

如果这位怀疑主义者还不相信他的感官，可以用照相机试试。照相机没有记忆，也不先入为主，故而只真实地记下实在的一切，但它会给我拍下同样的相片。在那边，它拍下全身像，缩短一半距离，它拍下身体的某部分，在这里，它没有把人拍下，什么也没有——否则便朝反方向拍，那就是宇宙。

* * *

因此这个脑袋不是一个脑袋，而是错置的观念。如果我在这里还能找到它，我就“看见了东西”，我就必须赶紧去找医生了。不管我发现的是一个人脑袋，还是一个驴脑袋，一个煎鸡蛋，或是一束美丽的鲜花，都无关紧要，因为不管带什么样的头饰，都会引起错觉。

但是，每当我清醒的时候，我肯定没有脑袋。另一方面，我不仅仅不是没有脑袋：实际上，我有好几个各有用处的脑袋。在人和照相机里，在电影画面上，在镜前刮脸时，从门的球形把手、调羹、咖啡壶以及其他一切铿亮的东西上，到处都有我的头像，尽管多少是扭曲、变形的，前后颠倒的，经常是倒过来的，而且是重重叠叠的。

但只有一处不会出现我的脑袋，这就是“我的双肩上”，这里会遮挡那中心虚空，而它正是我的生命之源。然而很幸运，无论什么都做不到这一点。事实上，这些游离的脑袋最多只能构成“外在”或表象的世界中那些不恒定和一般的偶然之事，这世界虽然与那中心本质并无二致，但不会给它带来丝毫的影响。我镜中的脑袋是最不具有特权的，因此，我没有必要将它作为我的脑袋：我很小的时候，就认不出镜子里的我，现在也是如此，我一时又获得了已经失去的天真。在我比较清醒的时候，我看见镜子后面那个房间里的人，那个很熟悉的人，看上去他整天把时间花于打量这屋子——那个瘦小的、呆板的、拘谨的、特别的、衰老的、十分敏感的人——这与我真正的自我没有丝毫相似之处。我从来就是这个没有年龄的、坚硬的、无量的、清晰的、无瑕的虚空：简直无法想象，我竟会将这具凝视的偶片同我明明白白地在此时此地并永远看到的自我混为一谈！

*　　　*　　　*

电影导演……是些讲究实际的人，他们对描述经历的再创造比之对发掘有这种经历的人的内心更感兴趣，但实际上这两者多少有些关系。举例来说，这些专家非常明白，在一部电影里，若是有辆由别人驾驶的车，我会无动于衷，而要是看到有辆很明显是由我自己驾驶的车，我就会欣喜若狂。在第一个例

子中，我是人行道上的观望者，看到两辆相差无几的汽车快速驶近并相撞，驾驶员被撞死，随后车辆也燃烧了起来，我觉得颇为有趣。在第二个例子中，我就是驾驶员，当然，像所有亲身驾车的人一样，没有脑袋，我的车是静止的（车在这里是多么的微不足道）。此刻，我的双膝在抖动，脚紧踏在加速器上，双手紧握方向盘，车的前罩被甩在一边，电线杆子向后闪去，道路像一条蛇七扭八歪，对面来了一辆车，起先是很小的一点，而后越来越大了，向我直冲过来，然后相撞，产生了巨大的火光，最后是死一般的寂静……我重重地倒在坐椅上，长长地松了一口气。我上了一个当。

这种亲自驾车的影片是如何拍摄的？一般有两种拍摄方法。第一种是拍一个没有脑袋的假人，摄影机放在脑袋的部位。另一种方法是拍一个真人，演员的脑袋向后或向一侧靠，以便在这个部位放置摄影机。换句话说，如果我自己要当这名演员，他的脑袋就必须让位；也就是说他必须是我这样的人。而我的一张有脑袋的全身照就完全不像了：它只是一个完全陌生的人的像片，只是错觉的结果。

人们居然要到拍广告片的人那儿去瞧一瞧有关自己的最深奥，也是最简单的真理，像电影这样的煞费苦心的现代作品居然会帮助人们摆脱连孩子和动物都没有的错觉，这真是咄咄怪事。但是别的时代也有同样的怪现象，我们人类自我欺骗的能力从来不曾完善过。对人类处境虽模糊但又深刻的认识，也许能够充分解释古代社会许多宗教信徒和传说为什么会广泛流传：游离身体、四处飘移的脑袋，独眼或无头的妖怪和幽灵，长着野兽脑袋的人的躯体，还有殉道者（像处于斩首极刑的查理一世）在他们的头颅被砍下之后仍然行走如常，话语不断。这些无疑是离奇古怪的想象，但它们同人类的常识相比，却更接近

此人的真实画像。

* * *

但如果我（与常识相违）真的无头、无脸、无眼，那么我怎么会看见你呢？长了一双眼又有什么用呢？实际上，“看”这个动词有两个完全相反的意思。当我观察到两个人互相交谈时，我们便认为他们互相**看见**了，尽管他们的脸没接触，相距有几英尺远。但当我看你时，你的脸就是一切，而我的脸则什么也不是。你是我的目标。然而（因此阻碍启蒙的是常识的语言），我们用这个小小的词来指代两种行为：当然，同样的词必须指同一事物！实际上，第三人称的事物或人本身之间的行为实质上是视觉交流，即连续的、自我包容的物理过程之链（包括光波、镜片、视网膜、大脑皮层的视觉区等），在这里，科学家找不到“心灵”或“视觉”的立足之地，即使是真有的话，也是大同小异。相反，真正的看是第一人称的，因此也不通过眼睛。拿圣人的话来说，只有佛、婆罗门、真主、上帝才能真正看见、听见或者有所体验。

D·R·霍夫施塔特

反　　思

哈丁为我们展示了一幅关于人类处境的有趣图画，是孩子气的和唯我论的。从理智的角度看，它使我们生气并大吃一惊：有谁真心地抱有这些观念而不觉得为难吗？但从我们内心

深处来说，它是非常清楚的。我们从内心深处无法接受自己死亡的观念。在许多人中间，这种想法已经埋藏了这么久，以至于我们忘记了个人的不存在这个观念是多么的不可思议。我们似乎能够从别人的死轻易地推断出将来有一天我们自己也不存在。然而，我死的那一天会是怎样的呢？无论怎么说，一天就是有光和声的一段时间；我死的时候，是无所谓光和声的。"啊！错了，怎么能说没有呢？"我心中有一个声音反驳道："不能因为我不在人间，不能感觉到它们，就认为它们不存在了！这种观点太富唯我论色彩了！"我心中的声音被一个简单的三段论法的威力所逼迫，已经不情愿地跨越了这样一个观念，即我是宇宙的一个必要的组成部分。这个三段论大致如下所示：

所有的人都是要死的

我是一个人

因此……我……是……要死的。

除了此处"我"代替了"苏格拉底"以外，这是一个最经典的三段论。有什么证据来证明这大、小两个前提呢？大前提假定一个抽象的范畴，即人的类，小前提假定我也属于人，尽管我和这一类中其他成员之间看上去存在着很明显的不同（哈丁一针见血地指出了这一点）。

类是概括的前提，它并不太令人吃惊，但是，超越部分天生技能形成类的能力则是一种相当发达的智力。看上去蜜蜂有"花"这样的类概念，但很难想象它们能够形成像"烟囱"或"人"这样的概念。狗和猫似乎能够进行一些新的分类，如"狗食盘子"、"门"、"玩具"等等。但是，人类在这方面独占鳌头，他们能制造出一个接一个的新范畴。这种能力是人性的核心，也是

欢乐的深刻源泉。体育广播员、科学家和艺术家创造了新的概念，它们汇入我们的思想词汇中给我们以极大的乐趣。

大前提的另一部分是死亡这个一般的概念。人们很早就发现万物会消失、毁灭。调羹里的食物吃完了，拨浪鼓从儿童的凳上摔下来了，妈妈出去了一会儿，气球爆了，报纸在炉火中化为灰烬，沿街的房屋被夷为平地等等。这些现象当然令人震惊不安，但它们仍能被接受。被拍死的苍蝇、被灭蚊剂毒死的蚊子——这些是建筑在先前的抽象概念之上的，我们由此便得出了死亡这一总概念。关于大前提就讲这些。

小前提十分棘手。我小时候，看到我身外的某些东西有相似之处，如外形、行为等，于是就形成了“人类”这个抽象的概念。然后这个特殊的类反过来“包围”我并吞噬我——人类认识发展到一定水平之后，必然会认识这一点。当时，这样的发现一定十分令人吃惊，尽管现在我们中的大部分人早已将这件事如何发生的情形忘得一干二净了。

尽管如此，最令人不知所措的一步是这两个前提的连接。在我们有能力形成这两个前提时，我们已经开始对简单逻辑的感染力肃然起敬。但是，这两个前提的突然结合却意想不到地给了我们一记耳光。这残忍、野蛮的一记耳光把我们打得双眼直冒金星——也许几天、几星期、几个月恢复不过来。实际上，竟会是几年，竟会是终身使我们难以恢复自制！但是我们还以某种方式制止了这场冲突，将它引向别处。

高级动物是不是有能力认识自己是属于某个类中的成员？一条狗是不是能够(无言地)思考这样的想法：“我发誓我看上去同那边那些狗一样？”想象一下下面这个血淋淋的场面。属同一类的20个动物围成一圈，一个残忍的人不停地在它们中间转来转去，最后走到那个厄运临头的动物面前，当着其他动物的面，

一刀将它刺死。可能其他动物都意识到了大祸临头，它们会不会这样想："那个死去的动物与我一模一样，我的鹅肉很快会像它的鹅肉一样被烹煮。啊，不！"

将自己与其他成员归为一类的能力看上去只有高级动物才有。一开始人们进行局部的观察："我有脚，你有脚；我有手，你有手；唔……"这些局部的观察能被归纳为一个总体的概念。很快，我由你长着一颗脑袋的事实推断出我也有一颗脑袋，尽管我看不见自己的这颗脑袋。但是，这走出自己的一步是很关键的，而且，从某些方面看来，是否定自己的一步。它同关于自己的许多直接经验相抵触。这有点像哈丁关于动词"看"的两种截然不同的意思。你"看"和我自己"看"是两件截然不同的事情。然而，这种语义上的差别一直被太多的局部观察所掩盖了。一开始我形成人这一类概念时，并没有考虑到我自己；而现在当我确认这种差别时又不加怀疑地把自己当作这个类中的一员。

因此，逻辑压倒了直觉。就像我们能最终相信地球是圆的——遥远的月球也是圆的——人居住在上面不会掉下那样，我们最后也会相信这种唯我论的观点简直是狂热病的表现。只有像哈丁在喜马拉雅山经历过的那种敏锐的觉察才会使我们重新考虑关于自己和他人这个古老的问题。这是有关意识、心灵和自己问题的根源。

我是不是有大脑？我是不是真的会死？在人生的旅途中我们都会时常思考这些问题。每一个富有想象力的人也许常常会认为，一切生活只是一个大玩笑或一场大骗局——也许是一个心理实验——被某个不可思议的超人所戏弄，他看着我们对那些显而易见的荒诞的东西到底能信得多深（属于这一类的信念有：我听不懂的声音会真有意义，有人能听肖邦的音乐或能吃巧克力冰淇淋，但可以不喜欢这些东西；光在任何参照坐标中

都是恒速运动；我是由无生命的原子所组成的，以及关于我自己的死这样的观念等等)。但是，很不幸(或很幸运)这个“协同理论”在挖自己的墙脚，因为它假设有另一个心灵的存在——事实上，假设存在着一个超理智因而不可思议的心灵——以此来解释其他的谜。

看起来我们只得低头服输，承认确有某些互不相容的性质存在。你可以自行选择。我们都在主观和客观的两个世界之间徘徊，而这种困惑对人性来说是很重要的。

3 心灵的再发现

哈罗德·J·莫罗维兹

一个世纪以来，在科学界一直有一个很奇怪的现象。而许多研究者却没有意识到这一点，其他人则拒不承认这一点，连对他们的同事也不承认这一点。然而，这种奇怪现象在蔓延。

事情是这样的。生物学家们曾经断言，人的心灵在自然的等级里有特殊的作用，可是现在他们却无情地去拥护作为十九世纪物理学特征的赤裸裸的唯物论。与此同时，物理学家们面对着不容置疑的实验结果，开始同严格的宇宙机械模式分道扬镳，认为人的心灵在所有物理现象中起着不可分割的作用。这个现象好比两门学科乘上了两列相对而驶的列车，各自都注意不到铁道另一边发生的事情。

由于心理学家与物理学家持有这两种截然不同的态度，使得心理学家左右为难。从生物学的角度看，心理学家研究的是与确定性的内核——亚微观的原子和分子世界没有丝毫关系的现象。从物理学的角度看，心理学家研究的是“心灵”，是一个

未经阐明的原始观念，这个观念似乎既带有根本性，又难以理解。很明显，这两个观点都包含一定的道理——若能解决这个难题，行为科学的基础就能变得更深广。

从各个层次对生命进行的研究，即从社会行为到分子活动的研究，在现代，是以还原论作为其主要的解释观念的。这种认识方法试图按照低级的，也许又是更带有根本性的概念来理解高一级的科学现象。在化学中，大规模的化学反应是通过观察分子的活动来解释的。同样，生理学家是通过研究微器官和亚细胞实体的活动过程来研究活细胞的活动的。地质学是利用结晶体的特征来描述矿物的形成和特性的。这些研究的本质就是通过根本的结构和活动来寻找答案。

卡尔·萨根的畅销书《伊甸园的飞龙》中的观点是心理学领域内的还原论的例证。他这样写道："我的大脑理论的最基本前提是，大脑的活动——我们有时也称之为心灵——是其解剖学和生理学的结果，仅此而已。"为了进一步指出这一思潮，我们发现在萨根的词汇里找不到像心灵、意识、知觉、认识或思想这样的字眼，只有突触、脑白质、蛋白质和电极这样的词汇。

像这类将人类行为还原为其生物基础的企图由来已久，最初始于早期达尔文主义者和他们研究生理心理学的同时代人。在十九世纪以前，笛卡儿哲学的核心——心身二元论——试图将人的心灵置于生物学的领域之外。随后，由于进化论者强调了人类的"猿性"，迫使我们用类似于研究灵长目甚至是其他动物的方法来进行生物学研究。巴甫洛夫学派强调了这个观点，使之成为许多行为科学理论的基石。尽管心理学家们就还原论究竟能走得多远这个问题并没有形成一致的意见，但是，许多人却都不假思索地承认人类行为可以分解为内分泌、神经和生理结构的活动。虽然萨根的前提不违背心理学的普遍传统，但是，它

却相当激进，因为它试图在最基本的层次上解释所有的现象。他曾说过“仅此而已”，我认为这句话是针对这一目标的。

在心理学各流派正忙于将心理学还原为生物学的同时，其他生命科学家也在寻找更基本层次上的解释。他们的观点可以在一位著名分子生物学代言人的著作中找到，他叫弗兰西斯·克里克。他的著作《分子和人》当时是对活力论的批判，这种活力论认为生物学是用物理学领域之外的生命力量来解释的理论。克里克写道：“生物学的现代运动的最终目标实际上是用物理学和化学来解释全部生物学。”他又说，他所谓物理学和化学指的是关于原子层次上的东西的知识，只有在这个层次上的认识才是可靠的。通过强调全部，他阐明了一种极端还原论的观点，这是在整个一代生化学家和分子生物学家中一直占主导地位的观点。

* * *

如果我们把心理学还原论和生物学还原论相结合，并假定它们部分重叠，那么，我们将得到一系列解释，其范围从心灵到解剖学和生理学，细胞生理学，分子生物学，直至原子物理学。所有这些认识都建立在对量子力学的定律，即关于原子结构和变化的最新、最彻底的理论的理解之上。在这个范围内，心理学成了物理学的一个分支，这使心理学家和物理学家都感到不安。

用物理学中的第一原则来解释人类所有现象这一企图，也不是什么新东西，早在十九世纪中期，欧洲的生理学家对此就已形成了十分明确的看法。这个学派的代表人物，埃米尔·杜布瓦-雷蒙，在1948年出版的一本关于动物电著作的前言中便表明了自己的极端观点。他写道，“只要我们有足够的办法，就有可能创立研究一般的生命过程的分析力学（牛顿物理学），甚

至可以解决自由意志这个问题。"

在这些早期学者的论述中，有一种过分的自信，后来托马斯·赫胥黎和他的同事们在为达尔文主义辩护时，也表现出这种自信。甚至在今天，我们在现代还原论者的理论中还可看到这种自信的痕迹。现代还原论者离开心灵而转向原子物理学的基本原则。这种自信在当代社会生物学家的著作中表现得最明显，这些人的观点给知识界带来了活力。不管怎么样，杜布瓦-雷蒙的观点同现代极端还原论者的观点是一致的，唯一不同的是基本的原则不是牛顿力学，而是量子力学。

就在心理学家和生物学家踏踏实实地一步步将心理学和生物学还原为物理学的时候，他们却没有发觉，物理学正在展示使他们的认识为之一新的前景。十九世纪末期，物理学所描绘的世界图画井井有条，事物根据牛顿力学和马克斯韦尔电学的方程式有条不紊地展现出来。这种过程无情地进行着，完全不受科学家的影响，因为他不过是一个旁观者。许多物理学家认为他们的研究对象在本质上是完整的。

而从1905年阿尔伯特·爱因斯坦创立了相对论以来，这幅井然有序的世界图景被不客气地搅乱了。这个新的理论假定，在相对运动的不同系统内的观察者所观察到的世界是不同的。科学家不再是袖手旁观者，而是变成他们所研究的系统的积极参与者了。

随着量子力学的不断发展，观察者的作用在物理学理论中愈来愈成为中心的问题，成为确定一个事件的重要因素。这样，观察者的心灵随之在理论结构内也成了不可缺少的考虑因素。这个正在日益发展的新理论极大地震动了早期的量子物理学家，并引导他们去研究认识论和科学哲学。据我所知，在科学史上著书立说的人还没有详述过他们的发现所包含的哲学和人文科

学的意义。

新物理学创始人之一维尔纳·海森堡，对哲学和人文科学问题十分关注，在《量子物理学的哲学问题》一书中，他谈到了物理学家必须抛弃认为存在着一个对所有观察者都适用的客观时间标准和在时空中存在着独立于我们观察能力的事件的看法。海森堡强调指出，自然法则不再研究基本粒子，而是研究我们关于这些粒子的认识——换句话说，是研究我们的心灵的内容。计算出了量子力学基本方程式的薛定谔在1958年曾写过一本题为《心灵和物质》的非凡的小册子。这本小册子由一系列论文组成，在这些论文中薛定谔从新物理学的结论转向一种神秘的宇宙观，他认为这种宇宙观相当于奥尔德斯·赫胥黎的"永久的哲学"。薛定谔是第一位赞同《奥义书》和东方哲学思想的量子理论家。现在有越来越多的著作阐述这种观点，其中包括两本广受欢迎的书，它们是弗里特乔夫·卡普拉的《物理学之道》和加里·朱卡夫的《舞蹈大师》。

量子理论所面临的难题在"谁杀了薛定谔的猫"这个著名的悖论里可以一目了然。在一个假想的状况中，一只小猫被放置在一个封闭的盒子里，盒里有一罐毒药和一把用来砸罐的杵锤。杵锤由一个记录仪控制，计录仪记录像放射性衰变这样的随意事件。实验时间的长短须足以使杵锤放松的概率达到1/2。量子力学代表一个数学系统，相当于一只活猫和一只死猫函数的总和，它们各有一半的概率。问题是，观察(记录结果)的活动是杀了还是救了那只猫。因为在实验者观察盒内情况之前，两种结果出现的可能性是对等的。

这个轻松的实验实质上却说明了一个十分棘手的概念上的难题。用更正式的话来说，就是描写一个复杂系统只能采用与一个实验的可能结果相关的概率分布。为了从众多的可能性中

选择一个，就需要一个标准。这个标准构成了与这种概率相区别的事件，因为这种概率只是数学的抽象。然而，物理学家根据这个标准所作的唯一简单、一致的描写还包括观察者对结果的意识。这样，物理事件与人的心灵内容是不可分割的。这种密切的联系迫使许多研究者以严肃的态度把意识作为物理结构的有机组成部分来考虑。这种认识将科学引向与实在论相对立的唯心论哲学观。

诺贝尔奖金获得者尤金·维格纳在他的题为“关于心-身问题”的论文中，概括了许多当代物理学家的观点。维格纳一开始就指出，许多物理学家已经重新回到思想——即心灵——是第一性的这一认识上。他接着又说：“若是不参照意识这个因素，要想得到完全一致的量子力学的定律是不可能的。”文章的最后说，对世界的科学研究导致将意识的内容作为最终的实在，这一结论是多么了不起。

在物理学另一领域内的新发现加强了维格纳的观点。信息理论和它在热力学中的应用导致了这样一个结论，即热力学中的基本概念“熵”是表明观察者没有认识到该系统内原子结构的测量标准。当我们测量一个物体的压力、体积和温度时，我们对物体内部的原子和分子的确切位置和运动速度都一无所知。那些我们不了解的信息量的数量价值同熵是成比例的。在早期热力学中，从工程意义上讲，熵代表了某一系统内那些无法用来作外功的能量。从现代观念来看，人的心灵又一次介入了，而熵不仅与该体系的状态有关，而且还与人类关于那个状态的认识有关。

现代原子论的奠基者们一开始并没有妄想将“心理主义”的图画强加于世界之上。相反，他们一开始也是抱着相反的观点，但为了解释实验结果，他们被迫退到了现在的地位。

现在，我们能够将心理学、生物学和物理学这三大学科的观点连成一片。将各类看法的代言人萨根、克里克和维格纳的态度结合起来，我们就会有一个相当出乎意料的整体图景。

首先，人的心灵，包括意识和反射思想，可以用中枢神经系统的活动来解释，这些活动反过来又可以还原为那个生理系统的生物结构和功能。其次，每个层次上的生物现象能够通过原子物理学，也就是说都可以通过碳、氮、氧原子的作用和相互作用来解释。第三也是最后一点，现在能够用量子物理学来作最全面解释的原子物理学必须将心灵作为该系统的基本因素。

这样，我们一步步地绕了一个认识论的圆圈——从心灵出发，最后又回到了心灵。这一连串推理的结果也许会给东方神秘主义者，而不是给神经生理学家和分子生物学家，提供援助和带来安慰。不过，这个封闭的圆圈是由这三门学科内知名专家的解释过程的直接结合造成的。由于个人几乎不可能采用几个这样的范例进行研究，因此，一般问题就很少有人注意了。

如果我们反对这个认识论的圆圈，就将面临两大敌对的阵营；一个是物理学，它宣称无所不包，因为它描述全部的自然；而另一个是心理学，它也包罗万象，因为它研究的是心灵，这是我们关于世界的认识的来源。面对这两种观点中的问题，我们也许还是回到那个圆圈上去的好，还是对它作更同情考虑的好。即使它让我们失去了绝对的东西，至少它还包含了心-身问题，并为我们提供了一个可供各种独立的学科相互交流的框架。这个圆圈的封闭状态为心理学理论家们开辟了最好的途径。

* * *

这种研究人类行为的、具有如此社会生物学特征的、严格的还原论方法，在更狭窄的生物学基础上也遇到了麻烦。因为它包含了从早期哺乳动物到人的进化过程是连续的这一假设，这就是说，心灵，或者叫作意识，并非是一种突变。但如果我们思考一下进化过程不连贯的戏剧性变化，就很难苟同于这个假设。宇宙本身的起源，即"大爆炸"，是宇宙中不连贯发展的例子。生命的起源，虽然同"大爆炸"相比不那么惊心动魄，但也不失为一个很好的例证。

由于发现了遗传分子中的信息编码，使统治宇宙的规律受到了极大的震动。比如，在遗传生命产生以前，我们可以暂不考虑气温或声音不稳定诸因素，这样，就能得出关于行星演变的十分精确的规律。然而，到了后来，发生在热声层次上的单分子事件就会造成极大的后果。这是因为，如果假设这个分子事件是自我复制系统内的转变，那么生物进化的整个过程就会被改变。一个单分子事件通过致癌甚至能够杀死一条大鲸，通过产生出某种毒害一个系统内的某个关键的物种的剧毒病毒来毁灭这个生态系统。生命的起源并没有否定物理学的基本规律，而是赋予了一个新特征：分子事件的大规模影响。这个规律的变化使进化历史无法确定，因此就出现了一个明显的断裂。

当代许多生物学家与心理学家认为，灵长目动物在进化过程中产生的反思行为也源于改变了规律的断裂。同样，这个新发现并没有否定生物学的基本规律，只是增加了一个使我们有必要从新的角度来重新考虑这个问题的新特点。进化生物学家劳伦斯·B·斯洛博金将这个新的特征说成是内省的自我想象。他着重指出，这个特征改变了人们对进化问题的看法，并且使得我们不可能把重大的历史事件作为进化规律中固有的原因。斯洛博金声称，这些规律都发生了变化，尽管人和哺乳动

物有相似的大脑生理结构，但不能拿适用于哺乳动物的规律来理解人。

许多人类学家、心理学家和生物学家都以某种形式对这个日益明朗的新发现作过讨论。这是经验数据的一部分，你不能为了保持还原论的纯洁性而将它束之高阁。这个断裂需要进行彻底研究和评价，但首先需要被人所承认。灵长目同其他动物有很大的区别，而人同其他灵长目也有很大的区别。

有些人对还原论不加鉴别，深信不疑，认为它能够解决心灵这个难题，对此我们发现还存在某些问题。我们在前面已经讨论了这个观点的缺陷。它不仅有缺陷，而且也很危险，因为我们对我们的同胞作出怎样的反应完全取决于我们在理论上如何看待他们。如果我们将同胞仅仅看成是动物或是机器，那么，我们人类之间绚丽多采的生活就会黯然失色。如果我们用研究动物群体的方法来寻找人类的行为规范，那么就会无视那些只有人类才会有的特征，若没有这些特征，人生就会变得枯燥无味。极端的还原论在道德规范领域内所作的贡献寥寥无几。此外，它为追求人本主义提供了一个错误的词汇库。

科学界在认识大脑方面已取得了明显的进展，我本人十分热衷于当代科学研究中有着重要地位的神经生物学。然而，我们不愿意放纵这种热情，不会听任它不顾科学而胡言乱语，将我们禁锢在哲学的立场上，这种立场否认人类最富有神奇色采的特征而使人性变得枯燥乏味。数个世纪以前，我们的还原论前辈为了庆祝科学摆脱了神学枷锁而付出了巨大的代价，这个代价就是低估反思存在和特征的重大意义，人的精神是科学的可观察数据的一部分。我们在相信精神的同时，仍可以不失为一名优秀的经验生物学家和心理学家。

D·R·霍夫施塔特

反　　思

《交叉小径的花园》是崔朋设想的一幅宇宙的图画，它没有完成，然而并非虚假。你的祖先同牛顿和叔本华不一样……他认为时间并不是绝对和统一的。他认为时间是一个无穷系列，是一个正在扩展着、正在变化着的分散、集中平行的时间网。这张时间之网包括了一切可能性，它的网线互相接近，交叉，隔断，或者几个世纪各不相干。我们并不存在于这种时间的大多数里；在某一些里，您存在，而我不存在；在另一些里，我存在，而您不存在；在再一些里，您和我都存在。在这一个时间里，我得到了一个好机缘，所以您来到了我的这所房子，而在另一个时间里，您穿过花园，发现我死了。在再一个时间里，我说了同样的话，但我却是个错误，是个幽魂。

——豪尔赫·路易斯·博尔赫斯
《交叉小径的花园》

现实仿佛在可能性的大洋中飘浮，从汪洋大海中有些可能性被选择出来成了现实。非决定论说，这样的可能性在某个地方存在着，并成为真理的一部分。

——威廉·詹姆斯

有人认为量子物理学中的神秘之处与意识中的神秘之处，总是有些相同的，这种看法十分引人瞩目。莫罗维兹在前面为我们描绘了一个认识论的圆圈，里面硬科学、美、怪诞和神秘主义都有恰当的量，“搭配合理”。但是，这种看法在许多方面都与本书一个重要的观点相对立，本书认为，从原则上讲，非量子力学计算机的心灵模式(包括一切与心灵有关的内容)是可能的。但是不管是对是错——现在就下结论说莫罗维兹的观点哪个对哪个不对尚为时过早——他的有些话还是值得我们深思的，因为，主观和客观意见的交流问题是量子力学中一个十分重要的概念难题，这当然是没有疑义的。特别是，按照一般的解释，量子力学给予某些被称为“观察者”的系统以特有的因果状态，但没有确切地说明观察者的情况(特别是没有说明意识究竟是否属于观察者状态的必要组成部分)。为了说清楚这个问题，我们必须对量子力学中的“测量问题”作一简单的概述，为了把问题说得更明白，我们打算用“量子水龙头”这个比喻。

假设一个水龙头有两个开关，一冷一热，你可以连续拧开每个龙头。水从龙头里流出来，但是，这样的龙头有个奇怪的特点：放出来的水要不都是热的，要不都是冷的，但不会是温的。这被称作为水的“温度本征态”。你要是想知道水处在哪种本征态中，只有伸手去试试水是热的还是冷的。实际上，在正统的量子力学中，情况还要更加复杂。正是你探手去试水温这个动作才将水置于这一种或另一种本征态之中。在这个动作发生之前，可以说水处在一种迭加状态之中(说得更确切一些，处在一种迭加的本征态之中)。

放出冷水来的可能性取决于开关的调整。当然，要是你只拧热水开关，流出来的只能是热水；同样，要是你只拧冷水开

关，流出来的只能是冷水。如果你将两个开关都打开，你就造成了一种迭加状态。如果你用一个调整系数反复试验，你就能够得出依照那个调整系数流出冷水的概率。然后，你可以改变调整系数，再进行试验。在某些交叉点上，流出热水和冷水的概率是一样的。这就像掷一枚硬币（这个量子水龙头还使人遗憾地记起许多浴室的淋浴龙头）。到最后你能够收集到足够的数据来画一张表，从表中可以看出作为开关调整系数函数的流出冷水的不同概率。

量子现象与此很相似。物理学家能够旋转开关，将种种系统置于迭加状态中，这种迭加状态与我们前面谈到的冷热水迭加状态十分相近。要是没有对系统进行过测试，物理学家就无法知道该系统正处于哪种本征态之中。当然，我们可以证明，从基本意义上讲，该系统本身不"知道"自己处于哪种本征态之中，我们只能说，只有当观察者好比用手"试水温"的那一瞬间，该系统才能随机地认定自己处于哪种本征态中。在观察的那一秒钟之前，该系统所进行的活动就好像它不处于本征态之中。就一切实际目的而言，也就一切理论目的而言，事实上是就一切目的而言，该系统不处于本征态之中。

可以想象一下，倘若你不伸手去试水温，那么，为了确定从量子水龙头流出来的水究竟是热的还是冷的，你就要进行无数遍的试验（当然，我们排除一切暗示的可能性，比如水蒸汽）。例如，你可以通过将水引入洗衣机来进行试验。但是，如果你想知道你的羊毛衫是不是缩水了，一定要等到打开机盖以后才能清楚（这是一个有意识的观察者所用的测试方法）。又如，用水龙头流出来的水来泡茶。同样，你非得等到亲口品尝之后才能知道这是不是冷茶。你也可以在水龙头下面放置一个测温记录器。同样，你一定要等亲眼看到记录器上的读数或标记之后，才能

知道水的温度。如果说你不能确定水的温度，那你也不能肯定读数是不是都已记在纸上了。这里有重要的一点需要指出，不管是羊毛衫，还是茶，还是测温器，它们本身都不是有意识的观察者，因此，它们也跟水一样听凭于水龙头，进入各自的迭加状态：缩水或不缩水、冷茶或热茶、读数是高还是低。

这些话给人这样一个印象：讨论的问题似乎与物理学本身无关，而只跟古老的哲学之谜有瓜葛，比方说，"要是树林中一个人也没有，一棵树倒地时会不会发出声音?"对于这些古老的谜，量子力学有新的见解。它认为，观察结果有两种，它们截然相反，一种是对迭加状态现实的观察结果，另一种是对表面混合状态而实际上总是本征态的观察结果，这种状态的特点仅在测量之时为观察者所揭示。总的说来，一股或是冷水或是热水的水流与一股总是冷水或总是热水的水流有很大区别，因为这两种可能性互相"干涉"，就像水上的波浪，彼此重叠(正如赛艇后面的尾浪，或是拿石片在水面上飘水片一样)。事实证明，这些干涉的后果具有统计的性质，所以只有当你洗了许多件羊毛衫和泡了许多杯茶以后，这种后果才会变得明显。有兴趣的读者可以在理查德·费因曼的《物理学定理的特征》一书中读到有关这一区别的颇具文采的论述。

薛定谔的猫所面临的困境进一步深化了这个观点，即在观察者干涉之前，连猫也能一直处在量子力学的迭加状态之中，也许有人会反驳道："慢着！一只活着的猫难道不跟人一样是一个有意识的观察者吗?"也许是。但是，请注意：这只猫也有可能是只死猫。要是只死猫，它就不会是有意识的观察者了。实际上在薛定谔的猫身上，我们创造了由两个本征态组成的迭加状态，其中一个有观察者身份，而另一个则没有！现在我们该怎么办？面临这样的处境使人不禁想起一个禅宗之谜(保罗·

雷普斯在《禅身禅骨》中对此有详细叙述)，这个谜是由一位禅宗大师提出的：

> 禅如一位在悬崖之上咬着树枝而悬空的人。他两手无依，双脚无靠，树下有一个人发问："菩提达摩为何从印度来到中国？"如果他不回答，那他就失败了；但如果真的作答，他就会从悬崖上摔下，粉身碎骨，那他该怎么办？

在许多物理学家看来，关于有观察者身份的系统与没有此身份的系统之间的区别是人为的，甚至前后矛盾。还有，说观察者的干涉会引起"波函数的崩溃"，也就是说，会造成突然跳入随机选择的纯本征态中，这个观点给自然界的最终规律带来了不稳定性。"上帝不玩掷骰子的游戏"是爱因斯坦一生的信念。

在量子力学里同时拯救连续性和决定论的做法人们称之为量子力学的"多元世界解释"。这是由休·埃弗里特在1957年首先提出的。根据这个十分奇特的理论，没有一个系统是非连续地跃入一种本征态的。迭加状态的进化是平稳的，它的各式分枝以平行的方式进行扩展。一旦有必要，迭加状态就会长出新的分枝来，而这些分枝就会带来各种新的可能性。比方说，在薛定谔的猫这个例子中，我们发现有两个分枝，它们是平行发展的。"那么，猫怎么样了？它自己觉得活着呢，还是死了？"我们会这样问，埃弗里特会这样回答："这要看你注意哪个分枝。在一个分枝上，它自觉活着，而在另一个分枝上，它就毫无感觉。"我们的直觉会提出异议，我们会问："这样的话，在那致命的分枝上的猫在死之前几分钟会有怎样的感觉呢？当然，一只猫不能同时有两种感觉！那么，在哪个分枝上蹲着真正的猫

呢？"

当你认识到此时此地这个理论对你意味着什么的时候，这个问题会变得更加棘手。对于你的一生中量子力学分枝上的每一点（这种点是成万上亿的）来说，你分裂成两个甚至是更多的你，坐在巨大的"普遍波"函数的平行但又分离的分枝上。在埃弗里特的文章中产生这个难题的关键地方，他十分平静地加了以下脚注：

> 在这时，我们遇到了语言上的困难。在观察之前，我们只有一个单独的观察者状态，而后来，对于观察者来说，他有好几个不同的状态，这些状态都是出现在一种迭加状态之中的。这些分离状态中的每一个对于一位观察者来说都是一个状态，这样，就有了不同状态所描述的不同的观察者。另一方面，与此有关的是同一个物质系统，从这个角度来看，也就是同一个观察者，这个观察者对于迭加状态的不同因素来说，处在不同的状态之中（也就是说，他在这种迭加状态的分离因素中，有着不同的经验）。这样的话，当我们想要强调与此有关的是一个单独的物质系统时，应该用单数；当我们想要强调在迭加状态的单独因素中的不同经验时，就应该用复数（例如，这样说，"观察者对量子A进行观察，之后，在由此而引起的迭加状态中的每个观察者们各自都感知到一个特征值"）。

埃弗里特说这些话时一本正经：那么，主观上的感觉又怎样呢？他没有讨论这个问题，只是轻描淡写地一笔带过，也许他认为这个问题没多大意义。

但是，我们不禁要问："那么，为什么我自己觉得我只生活在一个世界里呢？"这个么，按照埃弗里特的观点，你并没有这样的感觉，你会同时感觉到其他所有世界的存在，只是走向这个分枝的这个你才感觉不到其他所有的世界。这样说太令人目瞪口呆了。这不禁使我们想起本篇"反思"前两段生动的引文，它们使我们三思，最终的问题是这样的："为什么是这个我在这个分枝上呢？是什么使我——我指的是这个我——感觉到我自己并没有被分裂呢？"

一天黄昏，夕阳映照大海。你和一群朋友沿着潮湿的沙滩站成一线。海水轻轻地拍打着你们的脚，你们默默无声地观望着夕阳慢慢坠向地平线。你静静地瞧着，如痴似醉，你忽然发现阳光反射在一个个波峰上，形成由成千上万个橙红色的光点连成的一条直线——对准你的直线！"啊！我太幸运了，我碰巧正好站在这条直线上！"你这样想，"真是太遗憾了，并非人人都和我一样与太阳保持这等完美的统一。"此时此刻，你的每一个朋友的头脑里也闪过与你一模一样的想法……是一样的想法吗？

这种想法是"灵魂探索问题"的中心。为什么在这个身体里有这个灵魂（或者说为什么正好落在普遍波函数的这一棵分枝上）？既然有那么多的可能性，为什么这个心灵偏偏就跟这个身体长在一起呢？我的这个"我"为什么就不能属于另一个身体呢？要是你这样回答："你长在这个身体里是因为你的父母生育了这个身体。"这显然是循环论证，不能令人满意地解答这个问题。但是，为什么偏偏他们是我的父母，而不是另外两个人呢？假如我生在匈牙利，那谁会是我的父母呢？要是我是别人，那我会是什么样子呢？或者，要是别人是我，那又会是怎样呢？或者，我就是别人吗？我是其他任何人吗？是不是只有一个普遍的意识？觉得自己是一个单独的人，是一个个人，这是不是幻

觉？在我们自认为最可靠、最不出差错的科学的中心，这些令人不可思议的问题又重现了，真是有点让人费解。

但是，从某种意义上说，这个现象并不令人费解。在我们头脑中想象的世界与随着我们所体验到的这个世界一同进化的许多别的世界之间有明显的联系。有这样一个众所周知的年轻人，他把手中的雏菊撕成碎片，喃喃地说道："她爱我，她不爱我，她爱我，她不爱我。"在他的心目中，基于对心爱的人的两个不同的模式，（至少）有两个截然不同的世界。或者，关于心爱的姑娘只有一个心理模式，这相当于量子力学迭加状态的心理类似物，这样说是否会更准确一些？

当小说家同时用多种可能的方法构思时，他的人物不是（打个比方说）处在一个心理的迭加状态之中吗？如果小说始终没写成，分裂的小说人物也许会在它们作者的脑海里继续着各种悲欢离合的故事。如果有人提问，哪个构思才是真实的故事？所有的世界都是同样真实的。

同样，还有一个世界，即普遍波函数的一棵分枝，在这个世界里你不会犯这种使你十分后悔的愚蠢错误。你难道不嫉妒吗？然而你怎么会嫉妒自己呢？除了这个世界外，还存在另一个世界，你在那个世界里所犯的错误会更加愚蠢，并且你会十分嫉妒在此时此地这个世界里的这个你！

也许有一种方法，你可以将普遍波函数看作是高高在上的上帝大作家的心灵——或大脑，随你用哪一个词——这样，一切可能的分枝都同时并存。我们只是上帝大脑的子系统，我们也不再拥有特权，就像我们的星系也不是唯一真的星系一样。如果从这个角度看，上帝的大脑就像爱因斯坦一贯认为的那样，平稳地、确定地发展着。物理学家保罗·戴维斯在他的新作《其他世界》一书中谈及了这个问题。他写道："我们的意识沿着

宇宙永远发展的、分叉的道路，随机地走着一条路线，因此，正是我们，而不是上帝，在玩着掷骰子的游戏。"

但是，这样还是没有解开这个我们人人都要问一问的最根本的谜："为什么对自己的和谐的感觉偏偏沿着这个分枝，而不是沿着其他分枝传下去呢？我随机地选择了某一分枝，在这分枝上我可以感觉到自己在代代相传，这一选择的背后隐藏着什么规律？在我的自己分裂后，为什么我没有感觉到自己随着另一个我，沿着其他路线走下去呢？是什么把我系在这一刻沿着这一宇宙分枝进化的这个身体之上的？"

这些问题都是最最基本的，已经很难用明白的语言说清楚了。而且，从量子力学那里也得不到什么答案。事实上，埃弗里特轻描淡写地一笔带过的波函数，又重新在另一头冒出来，这次算是彻底崩溃了。它成了个人同一性的问题，而且比原先的那个问题还要难以捉摸。

硕大无比的普遍波函数有无数不断发展的分枝，在这些分枝上，并没有什么证据提供给量子力学或其他任何学说，没有什么埃弗里特，也没有什么量子力学的多元世界理论。要是认识到了这一点，你就会更深地坠入谜的深渊。在有些分枝上，博尔赫斯的故事并没有写成。甚至还有这样的分枝，在那里，此篇"反思"写得恰好与此处一样，只是结尾的那句话略有不同。

4 计算机和智能

A·M·图林

模拟游戏

我建议诸位考虑“机器是否能够思维”这个问题。首先，我们必须考虑如何给“机器”和“思维”下定义。所下的定义可尽量接近这两个词的通常意思，然而这种想法是很危险的。如果探究“机器”和“思维”这两个词的意义仅依赖于找出它们通常的词义，那么，就很难逃脱这样的结论，即对“机器是否能够思考”这个问题的意思和回答只能从诸如盖洛普民意测验之类的统计材料中去找，但这样做很荒唐。我不打算去想出一个定义，而是用另一个问题来代替这个问题，这个新问题与那个老问题紧密相联，并且相对来说又不那么模棱两可。

这个问题的新形式可用一种我们称之为“模拟游戏”的游戏来描述。参加游戏的有三个人，一个是男性(A)，一个为女姓

(B)，另一个为提问者(C)，性别不限。提问者坐在一间与前两人隔开的房间里。对于这位提问者来说，这项游戏的目的就在于识别那两个答问者的性别。他用X和Y代表两位答问者，这样，在游戏结束时，他可以说："X是A而Y是B，"或者说"X是B而Y是A"。提问者C可以向A和B提这样的问题，例如：

C：X请告诉我，他或她的头发有多长？

现在，假设X真的是A，那A就必须作答。A在此游戏中的目的在于让C作出错误的判断。因此，他就可以这样回答。

"我一头短发，最长的约9英寸。"

为了不使声音被提问者听出来，答案采用书面的形式，要是用打字机打出来，那就更好了。理想的条件是两个房间由一台电传打字电报机联系。如做不到，就由一个中间人传递问题和答案。而第三个参加者B的目的则是为了帮助提问者。她最好的策略也许就是吐真言。她可以在自己的答案后面补充这样的话，例如，"我是女性，别听他的！"但这并不管用，因为那位男性可以说同样的话。

我们现在设想这样一个问题："要是一台机器在这项游戏中扮演A的角色，那会是怎样的结果呢？"提问者在这样的游戏中是不是也会像一男一女进行游戏时一样老出错呢？这些问题就代替了我们原来的那个问题，即"机器是否能够思维？"

对新问题的批判

我们除了问"对这个新问题该怎样答覆？"也可以这样问："这个新问题是否值得探讨？"我们探讨后一个问题没有多少麻烦，这样就防止了无限扯皮的现象。

这个新问题有一个特点，它在人的体力和智力之间作了明确的区分。没有一个工程师或是化学家声称能够制造与人的皮肤不差丝毫的物质。也许在将来某一时刻，人真能创造这样的奇迹，但是，即便是这样，我们认为，试图给一台“思维机器”套上人造皮肤的做法也不会使这台机器更像人。我们所提问题的形式是在下述情况下反映这个事实的：提问者既看不见其他游戏者也听不见他们的声音。

这个标准的一些其他优点可通过以下示范问题和答案体现出来：

问：请就第4号桥，给我作一首十四行诗。

答：别让我干这个。我从不会作诗。

问：将34957和70764相加。

答：(停顿30秒钟后答道)105621。

问：你玩国际象棋吗？

答：玩的。

问：我将k置于k1，没别的棋子。你的k在k6、R在R1。现在轮到你走了。你该怎么走？

答：(停顿15秒钟后)走R-R8，将军！

这种问答形式似乎适用于所有我们想列入的人类活动。如果一台机器在选美中落选，我们决不追究它，同样，如果一个人在与一架飞机的赛跑中落后，我们也没什么可说的。这个游戏的条件规定这些缺陷是无关的。如果“证人”觉得可取，他们尽可以吹牛，将自己的魅力、力量和勇敢无畏吹得天花乱坠，但是，提问者不能够要求他们提供实际的证明。

这项游戏中有一个弱点，即形势对机器大大不利。要是这个人试图将自己伪装成机器，他会尽出洋相。他在做算术时，由于速度慢或不准确，立刻就会被淘汰。难道机器就不能从事

某件被认为应该是思维但又与人的思维有很大不同的活动吗？这个异议咄咄逼人，但是，我们至少可以这样说，尽管机器可以被造得能够十分满意地进行这种模拟游戏，我们也用不着为这个异议多费脑筋。

我们建议，在玩模拟游戏时，机器最好的策略是不要模仿人的行为。尽管有这种可能性，但我认为这不会有多大的效果。不管怎么说，在此我们丝毫也不想探讨游戏的理论问题，我们认为最好的策略是设法像人那样自然地回答问题。

该游戏中的机器

我们起初提的那个问题只有到了给"机器"这个词下了严格的定义之后才能看得更明确。我们十分自然地希望所有的工程技术都可以适用于我们的机器。我们也希望出现这样的可能性，即一位工程师或一组工程师制造一台能运转的机器，但是，机器的制造者们对它的运转情况无法进行满意的描述，因为他们使用的方法基本上是试验性的。最后，我们希望把以平常方式出生的人与机器相区别，要使一个定义满足这三个条件，是困难的。你也许会坚持，比如，一组工程师必须是同一性别的，但这很难奏效，因为还有可能由例如一个男人皮肤的一个单细胞培育出一个完整的个人。若能做到这一点，那是生物技术的一大创举，可庆可贺，尽管如此，我们不会因此就认为这是属于"制造一台思维机器"的例子。这使得我们放弃了允许使用一切技术的要求。我们将毫无顾虑地去这样做，因为我们注意到这样一个事实，即目前对于"思维机器"的兴趣都是由一种特殊的机器引起的，通常我们称之为"电子计算机"或者"数字计算机"。

如果我们接受这个建议，那么，我们只允许数字计算机参加我们的游戏……

数字计算机有一个突出的特点，它能够模仿任何其他离散机器，因此我们称它为万能机器。具有这种特征的机器产生了很深远的影响，若不考虑速度，就没有必要设计各种新机器去进行各种计算工作。这些计算工作可由一台数字计算机完成，只要配备了适于各种计算的程序。我们会看到，由于这个原因，从某种意义上说，所有数字计算机都是一样的。

关于主要问题的对立意见

现在我们可以说已经把前提搞清楚了，并准备就“机器能否思维？”这个问题继续进行辩论……我们不能全盘摒弃最初的那个问题形式，因为在考虑代换是否恰当这个问题时，会有不同的意见，因此，我们至少必须考虑这方面必须说些什么。

我想要是我先就这个问题阐述一下自己的看法，那么，读者就会觉得问题会更简单一点。首先让我们来看看这个问题的更确切的形式。我认为在50年的时间里，计算机的信息存储量可达到大约 10^9，这样的话，计算机在模拟游戏中就会很顺利，一般的提问者在5分钟的提问后，能准确鉴别的概率不会高于70%。原来那个“机器能够思维吗？”的问题我认为没有什么意义，不值得讨论。然而，我认为到本世纪末，由于词汇用法会有较大的变化，普遍的学术见解也会改变，那时候人们又能重谈思维机器而不会自相矛盾。我还认为，如果掩盖这些看法，决不会带来任何益处。人们普遍认为，科学家进行科学研究工作，总是从可靠的事实到可靠的事实，从来不受任何未经验证的假设

的影响，这种看法实际上是错的。假如能清楚地划分哪些是经过证实的事实，而哪些又是未经验证的假设，那也没有什么害处。假设往往是非常重要的，因为它们为研究暗示了有益的研究方向。

我现在谈谈与我的看法相对立的观点。

（1）来自神学的反对意见。思维是人的不朽灵魂的一种功能。上帝赋予每一个男人和女人以一颗不朽的灵魂，但从未将它赋予任何其他的动物或机器。因此，动物或者机器不能思维。

尽管我不能接受这种看法，但我试图用神学的语言来回答这个问题。如果将动物和人划为一个类别，我认为这个观点更有说服力；因为，在我看来，生物与非生物之间的差别远远要比人和其他动物之间的差别大得多。如果这种正统的观点出现在其他宗教社会里，它的主观武断性就更加明显了。伊斯兰教认为妇女没有灵魂，基督教对此有何感想？但是，现在暂不管这一点，让我们回到问题的焦点上来。在我看来，上面所引的论点对上帝威力的万能性有很大的限制。上帝对有些事情也是无能为力的，比如，无疑他不能让1等于2；但是，我们难道不应该相信，要是上帝觉得合适，他完全可以赋予一头大象以灵魂吗？我们可以希望，上帝可以通过自己的威力造成变种，由于变种，大象有了一个较发达的大脑，可以用来满足灵魂的需求。同一形式的论点也可以用来解释机器。只是看上去也许有点不一样，因为“轻信”就不那么容易了。但这其实只能说明，我们认为上帝不太可能觉得这些环境适合于授予灵魂。关于环境问题我们将在本文的其余部分进行讨论。在企图制造这样的机器时，我们不应该无礼地篡夺上帝创造灵魂的权力，就像不应该剥夺我们生儿育女的权力那样；在两种情况下，我们其实都是上帝意志的工具，为他所创造的灵魂提供住所。

然而，这仅仅是猜想而已。不管能用神学的论据来证明什么，我仍不以为然。过去已经证明这样的论据是漏洞百出的。在伽利略的时代，就有人提出，“太阳一动不动地悬着……整天都不慌不忙，不想落下”(《约书亚书》，10.13)以及“他为大地奠定基础，叫它永远不动摇”(《诗篇》，105.5)这些经文正好用来驳斥哥白尼的理论。从我们今天的知识来看，就觉得提出这种论据是徒劳的。过去没有这样的知识，情况便大不一样了。

（2）所谓“鸵鸟政策”式的异议。“机器思维后果太令人恐惧了。但愿机器永远也不会有思维。”

这种观点不如上面的说法那样直言不讳。但它对我们许多人都有影响。我们都倾向于认为人类在某个微妙的方面比其他生物要优越。要是能证明人一定是高一等的，那再好不过了，因为那样的话，他高居一切之上的地位就不会有危险了。神学的论点那样流行，很明显是与这种情绪密切有关。这种看法在知识分子中会更普遍，因为他们比其他人更尊重思维能力，因此也就对人类思维能力的优越性更加深信不疑。

我认为这个论点不怎么重要，不值得一驳。安慰一下也许更合乎情理；这种安慰也许能在灵魂轮回说中找到。

（3）来自数学的异议。在数理逻辑里有一些结论，可以用来证明离散状态的机器的能力有一定限度。这些结论中最著名的是哥德尔定理，此定理声称，在任何一个足够有力的逻辑系统里，都能够形成陈述，而所作陈述在本系统范围之内既不能被证明是对的，也不能被证明是错的，除非这个系统本身就不一致。丘奇、克利恩、罗瑟和图林等人还有别的在某些方面同哥德尔定理很相似的结论。图林的结论更容易考虑，因为它直接涉及机器，而其他人的结论相对来说是间接的：比方说，如要使用哥德尔定理，我们还需要某些附加手段，通过机器来描述

逻辑系统，而要描述机器还需要再通过逻辑系统。这个结论涉及一种机器，它实质上是一台万能的数字计算机。即使是这样一台机器，它对有些事情也是无能为力的。如果计算机被设计成能在模拟游戏中回答问题的话，那么对有些问题它是无法给予正确答覆的，而对另一些问题，不管你给它多长时间，它也答不上来。当然，尽管这台机器回答不了许多问题，但另一台机器却能给予满意的解答。我们现在只假定，对这些问题只要回答"是"或者"不是"就行了，不会出现像"你认为毕加索怎么样？"这类的问题。我们知道机器必定无法回答的问题是下述这类问题："这台机器有以下特点……这台机器会不会对任何问题作出'是'的回答？"这里省略的是对某台标准形式机器的描述……如果所描述的机器与那台被提问的机器具有某些相对简单的联系，那么，我们就能知道，答案不是错了，就是压根儿没有答案。这就是数学的结论，此结论认定机器能力有限，而人类智能则没有这种局限性。

如果想对这个论点作一简要的答覆，我们就能指出，尽管它已经证明任何一台特定的机器都是能力有限的，但它并没有任何证据说，人类智能就没有这种局限性。但我认为这个论点不能就这么轻易了结，每当其中一台机器遇到一个合适的问题，并作出了我们明知是错的回答时，我们无疑会产生一种优越感。这种优越感难道不会是错觉吗？这种感觉无疑是真实的，但我觉得这并没有多大意义。我们自己平时也经常对问题作出错误的回答，因此，就没有权利因机器犯了错误而幸灾乐祸。还有，我们对付一台机器当然易如反掌，但我们无法同时对付所有的机器而且不出差错。一句话，有可能人比一台特定的机器聪明，但也有可能别的机器更聪明，如此等等。

我认为，那些持数学异议的人大多数愿意接受模拟游戏作

为讨论的基点。而持前两种反对意见的人不大会对什么标准问题感兴趣。

（4）来自意识的论点。这个论点在杰斐逊教授1949年的李斯德演说中阐述得很明确，我摘引了其中的一段话："若要我们承认机器与大脑是一样的，除非机器能够因为感受了思想与感情而不是符号的偶然涂抹写出十四行诗或协奏曲来。也就是说，它不仅写了，而且也应知道自己确实这样做了。任何机器都感觉不到（不只是属于简易发明之类的人工信号）成功的喜悦，也不会因困难而沮丧，因受奉承而沾沾自喜，因犯错误而闷闷不乐，因见异性而神魂颠倒，也不会因欲望得不到满足而暴跳如雷或一蹶不振。"

这个论点看上去否定了我们测试的有效性。按照这种观点的最极端形式，你若要肯定一台机器是否能思维，唯一的途径就是成为那台机器并且去感受自己的思维活动。这样的话，他就能够向众人描述他自己的感觉，当然，没有人会知道这些话是不是实话。同样，依照这个观点，要想知道某人是否在思维，唯一途径就是成为那个特定的人。这实际是唯我论的观点。这也许是符合逻辑的，但若真是这样，那思想交流就太困难了。A会说："A在想，而B不在想。"而B会说："B在想，而A不在想。"我们犯不着为这个观点争执不休，我们不如客客气气地认为大家都在想。

我肯定杰斐逊教授不愿意采纳这样一个极端的唯我论的观点。他倒更有可能愿意把这个模拟游戏当作一个测试。模拟游戏（省略了游戏者B）在实际中经常采用"口试"形式，来鉴定某人是真的理解了，还仅仅是"鹦鹉学舌"，让我们看一看这种"口试"的情形：

提问者：你的十四行诗的第一行是这样的，"我欲比君为夏日，

如何？”要是将“夏日”改成“春日”，是不是也可以，或许会更好？

证人：这样一改便不合韵了。

提问者：改为“冬日”怎么样？这样也合韵。[①]

证人：是没问题。但有谁愿意将自己比作冬日呢？

提问者：你认为匹克威克先生会使你联想起圣诞节吗？

证人：在一定程度上，会的。

提问者：但是圣诞节是在冬天，我认为匹克威克先生对这个比喻不会在意。

证人：我想你也许在开玩笑。冬日的意思是指某一个典型的冬日，而不是像圣诞节那样特殊的一天。

这里不再赘引。如果那台写十四行诗的机器在这场“口试”中能够这样对答，杰斐逊教授会作何感想呢？我不知道他是不是会认为，机器只是在“机械地通过发信号”而作出这些答复；但是，如果这些答复如上面所引那样令人满意，前后一致，我认为他不会再把机器当作“一个属于简易发明之类的人工信号”。所谓“属于简易发明的人工信号”指的是一架机器里的一些设计功能，可以用来播放一个人念十四行诗的录音，只要拨动键钮你随时都可以听到这段录音。

综上所述，我认为大部分支持来自意识的异议的人都可以经过劝说而放弃原来的主张，不至于陷入唯我论的困境。这些人因此也就有可能愿意接受我们的测试。

我并不想给大家留下这样一个印象，即我认为意识并没有

① 原文用的是莎士比亚的十四行诗的第十八首，首句为"Shall I oompare thee to a summer's day?"改summer's day(夏日)为win er's day(冬日)，音节与韵都不变。——译者

什么神秘之处。比方说，要想确定意识在人体中的位置，就是一个谜。但是，如果我们没有解决这些谜，还是能够回答这个与本文有关的问题的。

(5) 来自种种能力限制的论证。这些论证一般是这样一种说法："我担保，你可以使机器干任何你刚才提到的事情，可你永远也不能使一台机器有x类的行为。"这类行为包括许多特征。我在这里援引一例：

> 要和蔼、机灵、美丽、友好……富于首创精神、富于幽默感、善于明辨是非、会犯错误……会坠入情网，喜欢草莓和奶油……能使别人钟情于它，通达世故……措词得当，长于反思……像人一样行为多姿多彩，富于创新……

说这些话一般都用不着证明。我认为这些话都是以科学归纳的原则为基础的。一个人在他一生中看到过成百上千台机器。他由所见所闻得出了一些普遍的结论。它们形态丑陋，应用范围狭窄，只要范围略有变动，它们就束手无策。此外，它们的行为方式也非常单一，等等，等等。他很自然地认为，这就是机器大体上的必备特征。大部分机器的能力限制与机器储存量太小有关(我在设想，储存量这个概念可以通过某种方式进行扩充，它不仅仅包括离散状态的机器，也包括其他机器。因为目前讨论还不需要讲究数学的准确性，所以定义准确问题并不要紧)。数年前，由于数字计算机在社会上还鲜为人知，要是你光说其特征而不提其构造，那么，人们就会以为你在信口开河。我想这也是因为人们使用了科学归纳原则的结果。当然，人们在使用这个原则时，大都是无意识的。一个小孩一朝被火烫，

十年怕烛台，我认为他这就是在使用科学归纳(当然，我也可以用许多别的方式来解释这一现象)。人类的行为和习惯看上去不适合运用科学归纳。如果你想获得可信赖的结果，你就要对时空的大部分进行研究。要不然的话，我们会(就像许多说英语的儿童那样)以为世界上所有的人都讲英语，再去学法语真是傻透了。

然而，关于我们刚才提到的许多能力限制，还要特别说几句，说机器没有能力享受草莓和奶油，这种说法会使读者觉得有点轻率。我们有可能使机器喜欢这些美味，但任何强迫这样做的企图都是愚蠢的。值得重视的是，这种能力限制对解释其他能力限制也有影响。比如，难以使人与机器之间形成那种像白人与白人之间，或是黑人与黑人之间的友好情感。

认为"机器不会出差错"这种想法有点令人费解。我们不禁要反问："它们出了错就更糟了吗?"让我们站在同情的立场上，看看这究竟是什么意思。我想我们可以用模拟游戏来解释这种说法。有人声称，在游戏中提问者可以向被试问几道算术题来分辨哪个是机器，哪个是人，因为机器在回答算术题时总是丝毫不差。这种说法未免太轻率了。(带模拟游戏程序的)机器并没有准备给算术题以正确的答案。它会故意算错，以蒙骗提问者。机器在做算术题时，由于对要出现什么样的错误作出了不妥当的决定，因而显示了机械的故障。我们对这种观点作这样的理解，其实也不太富有同情心。但是，我们对这个问题限于篇幅不可能进一步讨论。在我看来，这个观点的根源在于混淆了两个不同性质的错误。这两个错误我们称之为"功能错误"和"结论错误"。功能错误是由某些机械或电器故障引起的，这些故障导致机器不能够按照原指令工作。在进行哲学讨论时，我们很容易忽视发生这种错误的可能性；这样的话，我们实际上

是在谈论“抽象的机器”。而这些抽象的机器与其说是实在的物体倒不如说是数学的虚构。从定义上讲，它们不会犯功能上的错误。从这个意义上讲，我们完全可以这么说：“机器从不出差错。”当某一意思与来自机器的输出信号联在一起时，就会产生结论的错误。比方说，机器能够自动打出数学方程或英语句子。当机器打出一个错误的命题时，我们就认为这台机器犯了结论错误。很明显，找不到丝毫理由说，机器从不犯这类错误。一台机器有可能别的什么也不能做，只会连续打出“0＝1”。举这样一个例子也许太过分了，我们可以换一个例子：机器会想办法通过科学归纳来得出结论。这种办法有时无疑会导致错误的结果。

有人说，机器不能成为它自己思维的主体。如果我们能证明机器的某些思维是有某些主题的话，我们就能驳回这种说法。尽管如此，“机器活动的主题”确实有点意义，至少对于研究它的人来说是这样的。比如，如果一台机器试图解$x^2-40x-11=0$这个方程式，我们不禁会认为，这时，这个方程式本身就是机器主题的一部分。从这种意义上说，机器无疑能够成为它自己的主题。这对编排它自己的程序，对预测因本身结构变化带来的后果都会有所帮助。机器能够通过观察自己行为的结果，修改自己的程序，以便更有效地达到某种目的。这并不是乌托邦式的空想，而是不久的将来可能办到的事。

有人批评说，机器的行为比较单一。这也就是说，机器不能够有很大的存储能力。直到最近，达到1000数字的存储量都很罕见。

我们现在这里考虑的一些反对意见实际上都是来自意识的那个异议的改头换面。通常，如果我们坚持认为，一台机器有能力做其中的一件事，并对机器所能采用的方法进行描述，

那么，不会给别人多深印象。人们会认为机器所使用的方法(不管是什么方法，总是机械性的)实在是太低级了。请参见前面所引杰斐逊演讲中括号内的话。

(6) 洛夫莱斯夫人的异议。洛夫莱斯夫人的回忆录中曾对巴比奇的分析引擎作过详尽的记述。她这样写道："分析引擎没有任何意图要想创作什么东西。它能做我们知道该怎样去指挥它做的任何事。"(重点为她所加)哈特里引用了这段话，并补充道："这并不是说，就不可能制造能'独立思考'的电子设备，(用生物学的话说)在这种设备里，我们能够引起可以用来作为'学习'基础的条件反射。从最近的一些发展情况看，这种设想从原则上说是不是有可能，已经引起人们的极大兴趣和关注。但是，当时制造的那些机器并不具备这个特点。"

在这点上我完全同意哈特里的看法，我们会注意到，他并没有断言当时的机器还不具备这个特点，他倒是指出了，洛夫莱斯夫人所能获得的证明还不足以使她相信这些机器已具备了这个特点。从某种意义上讲，这些机器已具备了这个特点，这是极有可能的，因为，我们可以设想某些离散机器有这个特点。分析引擎实际上是一台万能数字计算机，因而，如果它的存储能力和速度达到一定水准，我们就能通过适当的程序使它模仿我们讨论的机器。也许伯爵夫人或巴比奇都没有想到这一点。无论怎么说，他们不必能提什么要求，便提什么要求。

洛夫莱斯夫人的异议还有另外一种说法，即机器"永远也不能创新"。这种说法可以用一句谚语"世上无新事"抵挡一阵。谁能保证，他的"独创性成就"就不是接受教育的结果，就不是因循著名的普遍原则的结果？这种异议还有另一个稍稍好一点的说法，即机器永远也不能"使我们惊奇"。这种说法有点直截了当，我能够针锋相对地加以反驳。机器经常令我吃惊。这主要

是由于我对机器能做什么估算甚少，更确切地说是由于即使我作了估算，也匆匆忙忙，马马虎虎。我也许这样对自己说："我认为此处的电压应与彼处相同；不管怎么样，就当是一样吧。"我自然经常出错，结果我大吃一惊，因为一俟实验完成，这些假设早被忘得一干二净了。我对自己的这些错误开诚布公，但在我证实了所惊讶的事时，人们也不会认为我在信口开河。

我的回答并不会使批评者就此缄口沉默，他也许会这样做，所谓大吃一惊都是因为我自己富于想象力的心理活动，与机器本身毫不相干。这样，我们重又回到来自意识的那个论证上去，而背离了吃惊不吃惊的话题。我们不得不认为这种论证方式是封闭式的，但是，也许值得一提的是，要将某物认作会使人惊奇的，则需要许多"富于想象力的心理活动"，不管这件令人吃惊的事件是由一个人、一本书、一台机器还是任何别的东西引起的。

我认为，那种认为机器不会令人吃惊的观点是由这样一个为哲学家和数学家特别容易犯的错误引起的。它是这样的一个假设，即心灵一接受了某个事实，由此事实所引起的一切后果就会同时涌入心灵。在许多情况下，这种假设十分有用，但是，人们会情不自禁地忘了这是个错误的假设。如果照这样做的话，其必然结果就是认为，仅仅从数据和普遍原则得出结论会毫无效力可言。

（7）来自神经系统连续性的论证。众所周知，神经系统不像离散状态的机器。若是关于撞击神经原的神经脉冲规模的信息有误，那么关于外冲脉冲的规模的信息就会有很大误差。既然这样，就可以论证：我们不可能用一个离散状态系统去模仿神经系统的行为。

离散状态机器同连续机器肯定不大一样，这一点没什么问题。但是，如果我们严格遵循模拟游戏的条件，提问者就不可能从这个差别得到任何有利之处。如果我们考察一下其他一些更简单的连续机器，有些话就能够说得更明白。一台微分分析机就足以胜任了(微分分析机是一种用作非离散状态计算的机器)。有些这样的机器能打出答案来，所以可以参加模拟游戏。一台数字计算机不可能猜准微分分析机对一个问题究竟作何答覆，但是，它倒是有能力给出正确回答的。比如，如果你要它回答π的值是多少(实际上约等于3.1416)，它就会在3.12，3.13，3.14，3.15，3.16之间作随机选择，其选择概率依次分别为(比方说)0.05，0.15，0.55，0.19，0.06。这样的话，提问者就很难分辨哪个是微分分析机，哪个是数字计算机。

(8) 来自行为变通性的论证。我们不可能总结出一套规则来囊括一个人在所有可想象的环境中的行为。比方说，我们可以有这样一条规则：行人见到红灯止步，见到绿灯行走，但是，由于某种错误，红绿灯同时亮了，那该怎么办？我们也许会这样决定，为安全起见最好止步。但是，这个决定还会有其他问题。要想总结出一套可以囊括一切人类行为的规则，哪怕是有关红绿灯的规则，看来都是不大可能的。对这些看法我全赞同。

从这一点可以得出，我们不能成为机器。我试图重新进行论证，但又恐怕做不好。似乎可以这么说："如果每一个人都有一套行动规则来制约他的生活，那么，人同机器就会相差无几了。但实际上不存在这种规则，因此，人不能成为机器。"这里，不周延的中项十分刺眼。我想谁也没有这样去论证过，但我相信实际上用的就是这样的论证。有人将"行为规则"和"行为规律"混为一谈，因此使这个问题有点模糊不清。所谓"行为规则"我

指的是像“见到红灯止步”这样的规则。对这类规则你能服从，并能意识到。所谓“行为规律”我指的是自然律，若用在人体上的话，就像“如果你拧他，他就会叫喊”这样的规律。如果我们将上面所引论证中的“制约他的生活的行为规律”改为“他用以制约自己生活的行为规律”，那么，这个论证中的不周延的中项就不再是不可克服的了，因为，我们相信，不仅是，用行为规律来制约生活意味着人就是某种机器(尽管这种机器并不一定就是一台离散状态的机器)，而且反过来说，如果是这样一台机器的话，那么，它就是由这种规律来制约的。然而我们很难像否认完整的行为规则那样轻易地否认完整的行为规律。只有通过科学观察，才能发现这种规律；无论在何种情况下，都不能说：“我们找得够久了，这种规律实际上不存在。”

我们可以找到有力的证据，证明这种说法是不正确的。因为，假定存在这种规律，我们肯定能够找到。对一台离散状态的机器来说，我们极有可能通过观察找到规律，预测其未来的行为，这要有一个合理的极限，比方说，在1000年之内。但事情并非如此。我曾在曼彻斯特计算机内输入了一个程序，这个程序用1000存储单元，这样，这个16位数的计算机就可以在两秒钟内作出回答。我绝对不相信任何人能仅从这些回答中充分了解这个程序，并预测对未试值的回答。

(9) 来自超感官知觉的论证。我想读者都很熟悉超感官知觉的说法，熟悉它的四种方式即心灵感应、千里眼、先知和精神运动的意义。这些令人不安的现象似乎在与一般的科学观念作对。我们多么想怀疑它们！不幸的是统计的证据使人至少对心灵感应不得不信。人们很难重新调整自己已有的观念以接受这些新事物，我们一旦接受了这些事物，就离相信鬼怪精灵为期不远了。朝这个方向前进的第一步是相信我们的身体除了按照

已知的物理学规律运动外，还按照未知的、但相近的规律运动。

这个论点在我看来十分有力。我们可以这样回答，许多科学理论尽管同超感知觉有冲突，但实际上还是可行的；事实上，我们要是对这些现象置之不理，依然能活得很好。这是一种甚为冷漠的安慰，恐怕思维这个现象与超感知觉有着特殊的联系。

基于超感知觉的更具体的论证大致如下："让我做一个模拟游戏，让一个善于接受心灵感应的人和一台数字计算机作为证人。提问者可以想象'我右手中的那张牌是哪个花色？'这样的问题。具有心灵感应或是千里眼的被问者在400张牌中可以答对130张。而机器只能随机猜对约104张，因此提问者就能作出正确的鉴定了。"这里有一个有趣的可能性。假使这台数字计算机有一个随机数字生成程序，那么，我们很自然就能用这个程序来决定给予什么回答。但是，这个随机数字生成程序又处在提问者的精神运动的能力作用范围之内。有了精神运动，计算机猜对的次数比概率计算还要高，因此，提问者就无法作出正确的鉴别了。而另一方面，提问者也能通过千里眼，不用提问就猜对。有了超感知觉，什么样的事都会发生。

如果允许心灵感应介入模拟游戏，我们就有必要严格规定测试方式。现在这情景就好比在模拟游戏中，提问者在自言自语，一个被问者正贴墙侧耳倾听。要是将被问者置入一间"防心灵感应室"，就能满足所有要求。

D·R·霍夫施塔特

D·C·丹尼特

反　思

我们在下面一段对话中将对这篇语言清晰、内容非凡的文章作出反应。很明显，图林愿意相信人同人创造的机器之间最终的区别，就在于人有超感知觉，而机器则没有。对此我们想说几句。如果大家从表面来理解我们的意见（而不将它当作考虑周密的玩笑），你们就会想知道我们为什么要提这样的意见。很明显，图林认为有很有力的证据来证明心灵感应的存在。然而，在1950年时情况确实是这样，而现在可今非昔比了。事实上，20年以后，这种论据已经不那样强了。自1950年以来，各种各样特异精神能力的声明层出不穷，而且经常由有名气的物理学家为之撑腰。某些物理学家到后来感到被愚弄了，并且收回了原先的公开声明，但不出一个月，他们又重新接受了另一种超自然的流行思想。但是，我们可以说大部分物理学家，乃至大部分擅长研究心灵的心理学家都怀疑有任何形式的超感知觉的存在。

图林感到"心安理得"是因为他相信超科学现象可以通过某一种方式同业已建立的科学理论相和解。在这一点上，我们同他的看法不一致。如果像心灵感应、先知和精神运动这样的超科学现象确实被证明是存在的，那么，对物理学定律来说就不仅仅只是修改的问题了，在我们的科学世界观内不来一场大革命是不可能解决问题的。人们也许会怀着激动的心情盼望这场革命早日发生——但是，激动中也夹杂着悲伤和困惑。对如此众多的事物如此适用的科学怎么可能是错的呢？这场从最基本的假设开始重新思考所有科学的挑战将是一次理智大探险，但是，许多年以来，我们需要这样做的证据却完全没有积累起来。

5 图林测试：咖啡馆里的对话

道格拉斯·R·霍夫施塔特

参加者

克里斯，物理系学生；帕特，生物系学生；桑迪，哲学系学生。

克里斯：桑迪，谢谢你向我推荐阿伦·图林的文章“计算机和智能”。这是篇好文章，它无疑引起我思考——思考我自己的思维。

桑迪：我很高兴。关于人工智能，你现在还和以前那样怀疑吗？

克里斯：你搞错了，我从来没有反对过人工智能。我觉得这玩艺儿很有趣——也许有点异想天开，但为什么不可以呢？我只是觉得，你们这些人工智能的拥护者低估了人的心灵，有些事情计算机永远也干不了。比方说，计算机能写出像普鲁斯特[1]的小说那样伟大的作品吗？丰富的想象力、复

[1] 普鲁斯特(1871—1922)，法国小说家。——译者

杂的人物性格……

桑迪：罗马不是在一天内建成的！

克里斯：从文章来看，图林这个人很有意思。他还活着吗？

桑迪：不，他在1954年就去世了，死的时候才41岁。他要是还活着的话，今年该67岁了。尽管他已成了一个传奇式的人物，很难想象他现在还活着。

克里斯：他怎么死的？

桑迪：差不多可以说是自杀。他是个同性恋者，因此不得不去对付来自外界的种种粗暴和愚蠢，最后，他受不了，就自杀了。

克里斯：真是不幸。

桑迪：的确令人伤心。最使我感到难过的是图林没能亲眼看到计算机技术和理论的惊人进展。

帕特：喂，你们是否打算告诉我图林文章讲了些什么？

桑迪：图林的文章实际上谈了两个问题。第一个问题是"机器能否思维？"或者说是"机器将能思维吗？"图林回答说："能！"他将否定的观点一一加以批驳，得出了机器能够思维的结论。图林试图加以说明的另一个问题是，前一个问题其实并没有多大意义。因为这个问题的感情涵义太多了。许多人一听到人就是机器，或机器也会思维时，就大惊失色。图林通过改变措词，从而避免了让这个问题带有过多的感情色彩。比方说，帕特，你觉得"思维机器"这个词怎么样？

帕特：说实话，我觉得这个词令人费解。你知道是什么使我费解吗？这是因为在报纸和电视的广告中老是在说"会思维的产品"或是"智能烤箱"等等、等等的。是信还是不信？真叫人为难。

桑迪：我也知道你说的那些广告，我看是把许多人搞糊涂了。

一方面有人告诫我们说，“计算机笨透了，你必须什么都给它们安排好”；而另一方面，广告又在天花乱坠地说什么“聪明的产品”。

克里斯：这当然是对的。不知道你是否听说这样一件事，一家计算机终端制造厂别出心裁，称自己生产的终端为“愚蠢的终端”。

桑迪：的确别出心裁，但这样做同样使人越来越糊涂。我一想到这一点，就立刻联想到“电脑”这个词。许多人囫囵吞枣地接受了这个词，而有的人连看都不看就将它拒之门外。但是，很少有人定下心来，对这个词进行认真的分析，看看它究竟有多少意义。

帕特：图林有没有提出过解决这个问题的建议，为机器设计一套智商测试方法？

桑迪：这很有意思，但机器还没有发展到能接受智商测试的地步。相反，图林倒是提出过一套测试方法，从理论上讲这个方法可以用来验定机器是否能够思维。

帕特：这种测试的结果是否能给出“是”或“不是”这样明确的答案？要是它的答案这么明确，我就有怀疑了。

桑迪：没有这样明确。从某个角度看，这正是它的优点。它表明“是”与“不是”之间的界限是模糊的，同时还表明整个问题十分微妙。

帕特：这样说来，同哲学中常见的那样，这只是一个措词问题。

桑迪：也许是这样，但是，这个问题的措词感情色彩强烈，因此，在我看来有必要探究这些问题并对至关重要的字眼加以认真研究。这对于我们认识自己具有重大意义，所以我们决不应该敷衍了事。

帕特：那请你告诉我图林测试是怎么回事。

桑迪：这个测试的基本思想是模拟游戏，这是他自己这么叫的。在这个游戏里，一个男子和一名妇女各自呆在一个房间里，提问者作为第三者用某种电传打字机向他们发问。提问者可以向双方提出问题，但提问时他并不知道谁在哪个房间里。对于提问者来说，他的目的就是确定那位妇女在哪个房间里。现在，那位妇女可以通过回答问题来千方百计地帮助提问者。但同时，那个男子也在千方百计地用女人的口气说话冒充妇女以迷惑提问者。如果他成功地蒙骗了提问者……

帕特：提问者只能看到书面的回答，是这样吗？而且假设回答者的性别可以在他或她的回答中鉴定出来？这个测试看上去确实有点意思。我非常想有一天能试一次。在测试开始之前，提问者是否认识其中一位回答者？回答者之间相互认识吗？

桑迪：这也许不太可能。如果提问者认识其中的一位，或两位都认识，那么在测试中什么样的潜意识暗示都可能出现。如果三个人互相都不认识，那么，测试的可靠系数就大得多了。

帕特：是不是什么样的问题都能问？没有什么限制吧？

桑迪：没有，丝毫没有。

帕特：这样的话，游戏难道不就局限于某些只与性别有关的问题了吗？不难想象，那位男士为了使提问者相信他是位女性，会竭尽全力去回答一些十分隐私的问题，而这些问题大部分妇女即使借着匿名的计算机联系也会觉得难以启齿相答。这样他的性别也就暴露了。

桑迪：听起来很有道理。

克里斯：还有这样一种可能，提问者为了鉴定哪位是女性，会

提一些关于传统的性别角色差异的细节方面的知识，比如像衣服的尺寸等等。这样，模拟游戏中的心理因素就会十分微妙。照我看，提问者的性别与游戏大有关系。很明显，妇女对于某些涉及个人隐私的问题要比男子敏感得多。

帕特：要真是这样的话，也许这就是区别男女的标准！

桑迪：晦……这倒是个新问题！其实我也不知道是否有人按照模拟游戏原先的设计认真地做过，要是用现代计算机终端来做的话，那就容易多了。坦白地说，连我自己也不敢肯定，不管试验的结果怎样，它能证明什么呢？

帕特：我也想到了这个问题。假设提问者是位妇女，她不能认定回答者中的哪一位是妇女，那么，这能证明什么呢？当然不能证明那位男子是一名妇女。

桑迪：一点没错！尽管我对图林测试一点也不怀疑，但我不明白模拟游戏究竟是什么意图，而图林测试是基于这种游戏之上的！

克里斯：模拟游戏是用来测验女性的，而图林测试是用来测验“思维机器”的，我对这两者都不以为然。

帕特：桑迪！从你的话中我得出这样的一个结论，图林测试只是模拟游戏的延伸，它只需要一台机器和一个人，各自关在一个房间里。

桑迪：没错。机器千方百计想让提问者相信它是一个人，而人又想方设法让提问者相信他或她不是一台计算机。

帕特：除了你说的“机器千方百计”这句别有用意的话之外，这些话是很有意思的。但是，你怎么知道这个测试会使人认识到思维的本质呢？也许它测试的东西与思维毫不相干。也许，随便举个例子，有人会这样说，如果一台机器舞跳得很好，让人认不出它是机器，这样你才能说它能够思维。

还会有人举出别的特征来。若是机器会愚弄人，叫他们坐在机前打出字来，这就那么了不起吗？

桑迪：我不能苟同你的看法。我以前也碰到过这样的反对意见，但说句老实话，我不明白。如果机器不会跳舞，不会掇起石块砸人，这又能说明什么呢？如果它能够与你无所不谈，这就证明它能够思维——至少对我来说是这样。我个人认为，图林一下子便清楚地在人类的思维与别的行为之间作了十分明确的区分。

帕特：现在是你叫人不明白了。如果我们并不能从那个能在模拟游戏中取胜的人身上得出什么结论，那么，一旦机器赢了图林测试，我们又能从它身上得出什么结论呢？

克里斯：这个问题提得好。

桑迪：如果那位男性被问者真赢了模拟游戏，我认为我们确实能够得出某个结论。当然，我们得出的结论不会是，他是位女性，而是他对女性心理有比较深刻的理解（如果真有女性心理这回事的话）。如果计算机玩花招，让人相信它是人的话，那么，我想我们也会得出类似的结论，即认为计算机对人是怎么回事，或者说对"人类状况"有比较深刻的理解（不管这种"人类状况"是怎么一回事）。

帕特：也许你有道理，但不能说这就是思维，对不对？我认为，通过图林测试只能证明被测机器成功地模拟了思维。

克里斯：我十分同意帕特的意见。众所周知，现在形形色色的计算机程序能够模拟各种各样的复杂现象。例如，在物理学领域内，计算机能够模拟粒子、原子、固体、液体、气体乃至星系的活动，但我们都很清楚，模拟毕竟是模拟，没有人会以假当真。

桑迪：哲学家丹尼尔·丹尼特写了一本题为《灵感》的书，在这本

书里，在谈及模拟飓风这个问题时，他也持同样的看法。

克里斯：对，这个例子很能说明问题。很明显，在计算机模拟飓风时，在它里面并没有真的飓风，因为它的存储并没有被每小时200海里的狂风撕成碎片，存放计算机的屋子里也没有发大水。

桑迪：噢，快别说了，你的论点不公平！首先，程序设计者并没有声明模拟就是真的飓风。它只不过是模拟了飓风的某些特征。其次，你说在模拟飓风里并没有时速为200海里的狂风，这个结论下得太快太早了。对我们来说，确实没有狂风暴雨，但是，如果这个程序设计得非常具体，在模拟时就会出现模拟的人，他们会像我们那样体会到飓风的袭击。在他们的心灵里，或者说在他们模拟的心灵里，这场飓风就不只是模拟了，而是真的自然现象，真的狂风暴雨。

克里斯：啊，好家伙，真像是一本科幻小说！好了，我们现在讨论的不是模拟单个的心灵了，而是模拟整个人类了！

桑迪：不，我的意思只是想说明，这种认为模拟的事物不是真实的事物的观点是错误的。观察者不能用一般的旧眼光去看待模拟现象，而应对模拟现象作出新的评估。事实上，他必须用一种特殊的观点才能辨认模拟的事物。这样的话，为了看见狂风暴雨，他必须戴上特制的"计算机眼镜"。

帕特："计算机眼镜"？我不懂这是什么意思！

桑迪：我的意思是说，你若想看见模拟飓风带来的狂风暴雨，你必须用恰当的方法来观察模拟。你——

克里斯：不，不，不！模拟飓风没有狂风暴雨！不管它在模拟的人看来有多大风雨，但它不会真有风雨！在模拟过程中，任何计算机都不会被飓风撕成碎片！

桑迪：当然不会，但你将几个不同的层次搞混起来了。物理学

定律也不会被真的飓风撕成碎片。在模拟飓风的过程中，如果你窥视计算机的存储系统，希望看到刮断的电线等等，会大失所望的。但是，如果在正确的层次上，也就是说在结构被编成密码储进存储的层次上，将会看到抽象的线路被刮断了，某些变项的值发生了很大变化以及其他一些情况。这就是你的狂风暴雨，这是真的狂风暴雨，只是有点隐而不露，不易察辨。

克里斯：对不起，这我无法接受。你认为我应该寻找一种新的风暴，这种风暴以前同飓风没有什么联系。按照你的看法，只要通过你的特殊"眼镜"就可以看到某些可称作"狂风暴雨"的后果，那么，随便什么东西都可以叫作飓风了。

桑迪：太好了。你的理解一点不错！你要认识一切飓风，你必须先认识飓风带来的后果。你不可能进入其中，在风眼正中找到虚无缥缈的"飓风的本质"或是"飓风的灵魂"之类的东西。由于存在着某种模式——一种有一个风眼及其他特征的旋转风暴，才使你认为它是一种飓风。当然，你可以说一场飓风还有许多其他的特征。

帕特：气候现象是一个至关重要的先决条件，你难道不这样认为吗？计算机里面的东西怎么可能成为一场风暴呢？在我看来，模拟不过是模拟而已。

桑迪：那么，我想你会认为计算机进行的计算也是模拟的，也是假的。只有人才能进行真正的计算，对不对？

帕特：计算机得出正确的计算结果，所以它们进行的计算并不完全是假的，但它们不过是些模式。绝对不存在理解的问题。就拿点钞机来说吧。说句老实话，你认为点钞机在清点钞票时会感觉到自己在进行计算吗？计算机只不过比点钞机更先进、更复杂而已，这是我对它的看法。

桑迪：你说点钞机在工作时不可能像小学生做算术时那样感觉到自己在进行计算，这一点我没有意见，但是，难道这就是“计算”的含义吗？这是不是计算的一个不可分割的部分呢？如果是的话，同我们大家的想法恰恰相反，为了进行真正的计算，我们不得不设计异常复杂的程序。当然，这这个程序有时也会疏忽，也会犯错误，有时会字迹潦草，令人难以辨认，有时也会在纸上乱写乱画……它并不比邮局职员用手计算金额可靠多少。但我相信这样一个程序最终会编出来的。这样，你对邮局和小学生如何进行计算就会有所理解啦。

帕特：我不相信你能编出这样的程序！

桑迪：也许能，也许不能，但这不是我的意思，你说点钞机不能进行计算。这使我想起了丹尼特的《灵感》中另外一段我十分喜爱的话，这段话有点讽刺性，这正是我喜爱它的原因。这段话是：“点钞机并不能进行真正的计算，它们只会转动齿轮。但是点钞机实际上也不能真的转动齿轮，它们只会遵守物理学定律。”丹尼特的原话是针对计算机说的，我把计算机改成了点钞机。同样，你也可以拿这段话去套人：“人并不能进行真正的计算，他们只能操纵心理符号。但是他们实际上并不能真的操纵心理符号，他们只能激活各式各样的神经原。但实际上他们并没有能力使神经原激活起来，他们只能利用物理学定律来激活神经原。”等等，等等。难道你没有发现，丹尼特这种独创的归谬法会让你得出这样一个结论，即计算实际上是不存在的，飓风是不存在的；一句话，高于粒子和物理学定律的任何东西都是不存在的？说计算机只不过拨弄符号，并没有真做计算，又能说明什么呢？

帕特：你举的这个例子也许有点过分。但是，它将我的观点解释得很明白，我认为真实的现象和模拟之间有很大的距离。飓风是这样，人的思维也是这样。

桑迪：唔，我不想再这么争论下去了，我再给你举个例子。假设你是个无线电收发报业余爱好者，你正在听另一个人用摩尔斯密码与你通话，然后你也用同样的密码向他回话，当提到"线的另一端的那个人"这样的话时，听上去不觉得很滑稽吗？

帕特：不，不滑稽，我觉得这话很正常。当然，线的另一端是不是真有一个人只是个假设而已。

桑迪：你说得对，但是，对于这种假设你不太会去证实的。你总是通过这些不平常的渠道认为对方是有人的。你既不必看到人的影子，也不必听到人的声音——你只需一种相当抽象的表示——密码。我的意思是这样的。你要是想通过密码的点点划划"看见"发报的人，你就得做一些译码的工作，一些翻译工作。这不是直接的感知，而是间接的。你得剥去一二层之后才能发现其隐藏在底下的实在。你必须戴上你的"无线电收发报业余爱好者眼镜"，才能"看见"密码信号后面的那个人。同模拟飓风一模一样！你并没有看见放置计算机的屋子里乌云笼罩，你不得不破译计算机的存储。你不得不戴上特殊的"存储破译眼镜"，然后，你看见的才是一场飓风。

帕特：呷，哦，哦！你说得天花乱坠了，等一下！就拿短波无线电来说吧，是真的有人在那边，也许在斐济岛上，或别的什么地方，我坐在无线电旁破译密码，就说明那个人的存在，这如同你看到一个影子就会得出结论说，必定有一个造成影子的物体那样。当然，我们决不会将影子和物体

混为一谈！让我们再回到飓风这个例子上来。在模拟飓风的背后，并没有真正的飓风，计算机进行模拟并没有依照一个飓风原形。实际上，你看到的是一个飓风影子，并没有真正的飓风。我不愿把影子和现实混为一谈。

桑迪：好吧。我不想说得那么死。我甚至还承认，模拟飓风就是飓风的观点是很愚蠢的。但我想指出，这个观点并不像你乍听起来那样愚蠢。一旦你考虑模拟思维，你手头的东西就跟飓风大不一样了。

帕特：我不明白。灵感在我看来就像一场心理的飓风。但说正经的，你先得说服我才行啊。

桑迪：好吧，为了说服你，我先还要就飓风再补充两点。

帕特：哦，不！行，说吧，说吧。

桑迪：谁也说不清楚，就是说不能用非常确切的话来说清楚，飓风究竟是什么样子。许多风暴都有一些共同的抽象的特征，正因为如此，我们把这些风暴称为飓风。但要在飓风和非飓风之间划一条明显的界限，那是不大可能的。因为形形色色的风暴太多，如旋风、暴风、台风、尘卷风等等……木星上的大红点算不算飓风？太阳斑点是不是飓风？风洞里有没有可能出现飓风？在试管里会不会出现飓风？你完全可以借想象，把中子星表面的微观风暴也当作是飓风。

克里斯：那也不是牵强附会的，你知道“地震”这个概念现在实际上包括了中子星上发生的震动。天体物理学家认为，脉冲星上的脉动现象是内震动引起的，可以称之为“星震”；这正是中子星表面发生的震动。

桑迪：对，记起来了。“星震”在我看来虽有点光怪陆离，却十分令人神往。这是发生在超现实表面上的超现实震颤。

克里斯：你能想象在一个巨大的、旋转着的由纯核物质构成的球体上发生的板块构造吗？

桑迪：这个想法确实十分大胆。因此星震和地震可以纳入一个新的、更抽象的范畴之中。科学就是这样，不断地扩大大家所熟悉的概念，但又使它们与熟悉的经验之间的距离越来越远，当然，其中最根本的东西是不会被改变的。数的系统就是这样一个典型的例子。数的概念从正数扩大到负数，又扩大到有理数、实数和复数；瑟斯博士说过，这种扩大过程是不会有尽头的。

帕特：桑迪，我想我明白你的意思了。我是学生物的，在生物学中，许多亲属关系都是通过抽象的方法加以确立的。在确定某种物种究竟属于哪一科时，我们是根据这些物种在某一层次上共有的某个抽象特征而定的。如果你的分类是根据非常抽象的模式来进行的话，那么，有许多现象可以被认为同属"一类"，尽管在这一类之中的许多现象在表面上互不相近。因此，我对你的观点略有领悟，你无非是说，模拟飓风从某种我觉得很可笑的意义上说，可以是一场真的飓风。

克里斯：也许你想说的不是"飓风"，而是那个"是"吧！

帕特：为什么？

克里斯：图林可以引申"思维"这个词，我为什么就不能引申"是"呢？我的意思无非是说，当有人蓄意将模拟事物与原事物混为一谈时，他实际上是在进行哲学上的诡辩，借以骗人。这比单单引申几个名词含义更严重。

桑迪：你说"是"的概念扩大了，这一点我觉得很有道理，但是，诡辩、欺骗之类的话太过头了。不管怎么说，如果你不反对的话，我再就模拟飓风说件事，然后我们就可以讨论模

拟心灵了。假如有这样一个模拟得很逼真的飓风，也就是说它模拟了每一个原子。当然，我承认不可能这么逼真。这样，这个模拟飓风就会拥有确定“飓风本质”的一切抽象结构，我希望你会承认这一点。那么，究竟是什么原因使得你不能叫它为飓风呢？

帕特：我看你是在撤回相等的要求！

桑迪：没错。可是这些例子摆在那里，我被迫回到我原先提的要求上。我说过我会食言的。现在让我们看看思维，它才是真正需要考虑的问题。思维比飓风更抽象，它是描述发生在大脑这个媒介中的一些复杂事件的方式。但实际上，世界上有几十亿个大脑，而每一个大脑里都发生着思维。也许每一个大脑的生理状况都不尽相同，但是，在里面进行的都是“同样的活动”，那就是思维。那样的话，媒介本身并不重要，重要的是抽象模式。每一个大脑中都有同样的活动，因此谁也不能说他的思维活动就跟其他人的不一样。现在，让我们假设同样类型的活动也会发生在某种新的媒介内，你难道能否认说在那里没有思维吗？

帕特：这也许很难否认。但是，你换了一个问题。现在的问题是这样的：你怎样才能确定在那里面真的发生了“同样类型”的活动？

桑迪：图林测试的妙处就在于它告诉你在什么时候可以确定！

克里斯：我一点也看不出。你怎么知道在计算机和我的心灵里进行的活动具有同样的类型呢？仅仅是因为它同我一样回答问题吗？你所注意的都是一些外在的东西。

桑迪：在我对你说话时，你怎么知道在我内心有类似你称作“思维”的活动在进行呢？图林测试是个异想天开的探索，有点像物理学中的粒子加速器。克里斯，我想你会喜欢这

个比喻的。在物理学中，如果你想要了解在原子或亚原子层次上的情况，由于无法直接观察，你可以将加速的粒子散射在靶上，以便观察它们的行踪。通过各种方式，你可以推断出靶的内部本质。图林测试将这个方法用于研究心灵。他把心灵看作是无法直接观察到的"靶"，但是，这个"靶"的结构可以更抽象地推知。通过将一些问题"散射"到一个心灵之靶上，你可以了解这个靶的内部结构，就跟物理学中的情况一样。

克里斯：说得更准确一些，你可以假设这样或那样的内部结构来说明所观察到的行踪。但你所假设的内部结构事实上可能存在，也可能不存在。

桑迪：请等一等！你是不是在说，原子核只是假设的实体？它们的存在——或者说"假设的存在"——毕竟是由原子散射的粒子的行踪所证明，或者说是由它们所"暗示"的。

克里斯：我觉得，物理系统一般说来要比心灵简单得多，因此，推理的可靠性也就相应地更大些。

桑迪：做实验和分析实验的难度也相应增大了。在图林测试里，仅在一小时之内你就可以做许多相当复杂的实验。我认为人们以为彼此都是有意识的，这仅仅是因为他们不停地相互监察。这件事本身就很说明问题。

帕特：大致是对的，但这比人们用电传形式进行对话复杂一些。因为我们亲眼看见了别人有身体，有脸，有表情。他们是我们的同类，因此我们觉得他们有思维。

桑迪：在我看来，这完全是人类中心主义的思维观。那是不是说，你以为橱窗里的时装模特儿因为更像人，所以是有思维的，而设计巧妙的计算机则没有思维呢？

帕特：显然，要给一个实体以思维能力，光外表像人是不够的。

然而，机体性质和种族的异同无疑会达到一定的可信性，这是十分重要的。

桑迪：我不同意你的这种说法。这简直是十足的沙文主义。我觉得关键的问题在于内部结构的相似性。不是在身体、机能和化学结构上的相似性，而是在组织结构，也就是软件方面的相似性。在我看来，一个实体是否有思维实际是这个实体的组织是否能通过某种方式进行描述的问题。而我绝对相信，图林测试的目的就是为了探明某个实体是否具有这种组织。你单凭我的躯体状况认为我有思维，我觉得这种看法很肤浅。在我看来，图林测试实际比单纯的外在形式要深刻得多。

帕特：你并没有完全使我信服。并不是仅仅有身体形状才使我觉得里面有真的思维。我已经说过了，还有种族的异同这个考虑。我和你都是由DNA分子创造出来的，我觉得这就十分重要。这样说，人体的外形表明人体具有同样的生物学的渊源，正是因为这个原因，人们才认为具有这样身体的人是有思维的。

桑迪：但这一切都是间接的证据。你肯定希望得到一些直接的证据。图林测试所寻找的正是直接证据。我认为这是测验思维唯一的办法。

克里斯：但你会被图林测试所愚弄，就像在测试中提问者会错误地认为那个男性被问者是位妇女一样。

桑迪：这一点我承认。如果这个测试进行得太快、太肤浅的话，我有可能上当受骗。但我所要寻找的是我所能想到的那些最深刻的东西。

克里斯：我不知道设计的程序是否能够理解玩笑。我认为这是测试智力最好的办法。

桑迪：你的意思是说，幽默对于一种据认为有智力的程序来说，无疑是一场严峻的考验，这一点我同意。但还有一个方面也同样重要，也许还更加重要，这就是测试程序的感情反映。因此，我想问问它对某些音乐作品和文学著作有何感想，尤其是那些我自己最喜爱的作品。

克里斯：要是它回答你说“我不知道这个作品”或者“我对音乐没兴趣”，那你会怎么想呢？如果它没有任何的感情反映，你该怎么办？

桑迪：这就有问题了。如果它对某些话题避而不谈，那么，我就会怀疑它是不是有思维。

克里斯：你为什么要这样说呢？你为什么不说，跟你打交道的那个玩艺儿只有思维而没有感情呢？

桑迪：你这话说在点子上了。只是我坚信感情和思维是不能分开的。换句话说，我相信感情是思维能力必然的副产品。思维的本质是感情的基础。

克里斯：嗯，但假如你错了呢？如果我设计了一台有思维但没有感情的机器，你会怎么想呢？那么，我的那台机器的智力会得到你的认可，因为它通不过你的那种测试。

桑迪：我希望你向我指出，感情问题与非感情问题的界限在哪里？你也许会问一部伟大的小说有何重大意义。这个问题就要求对人类感情的理解！这算是思维呢，还只算是冷静的计算？你也许会询问如何选择某些微妙的措词。要想回答这个问题，你就需要了解这些词的涵义。图林在他的那篇文章里列举了类似的例子。你也许会问机器怎样处理复杂的男女关系，它需要大量了解人的动机及根源。如果机器通不过这样的测试，我就很难说它有思维了。就我个人的看法而言，思维能力、感觉能力和意识都是一个现象的

三个不同侧面，三者相互依赖，缺一不可。

克里斯：你为什么就不能设计这样一台机器，它没有任何感觉，但是能够思维、能够作出复杂的决策？我觉得这并不矛盾。

桑迪：但我觉得有矛盾。你在说这句话时，脑海里浮现出来的是一台金属的、有棱有角的机器，也许放置在一间带有空调的房间里。这台机器摸上去是硬的，方的，冷的，里面装着千万条彩色电线，它静静地躺在地板上，"嗡嗡"地响，打出字条来。这样一台机器也许是个对弈的好手，我开诚布公地说，它能作出许多决策。但是，我决不会认为，这台机器因此就有意识了。

克里斯：这话怎么讲？在机械论者看来，能对弈的机器已具有思维的雏形。

桑迪：我认为，意识来源于一个十分周密的组织结构，而对于这样一个组织结构我们还不能作详尽的描述。但是，我相信我们会逐步理解它的。所谓意识，它必须能够从内部反映外部世界，并且能够从内部的反映模式出发，作出对外部存在的反应。还有一点补充，一台意识机器必须具备这样一个重要的特点，即它能够包含一个相当发达的、灵活的自我模式。正是在这个关节点上，包括最好的对弈机在内的所有程序都逊色了。

克里斯：对弈程序不是也能预见吗？它会对自己说，下一步棋怎么走。它会这样说："要是你走这儿，我就走那儿；如果你这样走，我就那样走……"这难道不是一种自我模式吗？

桑迪：并不真是。要是你非得说它是的话，那也只是一个局限性很大的自我模式。这是对自我最狭隘的理解。比方说，一个对弈程序决不会明白它为什么在对弈，并且不知道它自己

只不过是一个程序，也不知道它自己是在计算机内，更不知道它的对手是个人，它从未尝过赢输是种什么样的滋味。它也不知道……

帕特：你怎么知道它没有这些感觉呢？你怎么可能假定一个对弈程序有什么样的感觉或认识呢？

桑迪：噢，当然知道！我们都明白有些东西是决不会有感觉或认识的。你扔一块石头，这块石头当然不会知道抛物线是怎么回事。你摇一把芭蕉扇，这把扇子对空气的特性当然一窍不通，你说我不能证明这些命题，这是对的，但这里我们已经涉及到单纯的信仰问题了。

帕特：你的话使我想起一个我读过的道家故事。这个故事是这样说的。从前有两个圣人站在一座小桥上。一个说："我希望自己变成一条鱼。你瞧鱼在水中多么快活！"另一个说："你怎么知道鱼在水中是不是快活呢？你又不是鱼。"这一个的回答是："可你也不是我，那么，你怎么知道我是不是了解鱼的感觉呢？"

桑迪：这个故事说得太棒了！我们在讨论意识这个问题时，需要有严谨的态度。不然的话，你就会不是误入唯我论的歧途——我是宇宙之内唯一有意识的生命，就是误入泛心论的歧途——"宇宙万物皆有意识！"

帕特：那你是怎么知道的？也许宇宙万物是有意识的。

桑迪：有些人声称，石头，甚至像电子这样的粒子也是有某种意识的，要是你同这些人同流合污，我想我们便从此分手吧。这是神秘主义的论调，我可无法理解。至于说到对弈程序，我碰巧对它们了如指掌，我可以肯定地告诉你它们是没有意识的！绝对不可能有！

帕特：为什么没有呢？

桑迪：在它们里面只有关于象棋的最基本的知识。对对弈机来说，所谓“对弈”，就是将众多的数作比较，然后选择其中最大的一个数的重复的机械动作。对弈程序没有感觉，输了不觉丢脸，赢了也不觉光彩。它的自我模式还十分原始。它从不会自发地多花一点力气，它只要能下棋就算完成任务了。但是，有一点很有意思，我们还会说什么对弈机的“愿望”。我们说：“它想把王放在一排兵的后面，”或者“它喜欢早点把车移出来”，或者“它说我没有看见这着险棋”。

帕特：我们在谈论昆虫时也是这样说的。比方说，我们看见一只孤独的蚂蚁，就说：“它想回家”，或者“它希望把这只死蜜蜂拖回蚁窝去”。事实上，我们对任何动物都使用具有感情色彩的词汇，但我们实际上并不知道它究竟是怎么感觉的。我们说狗和猫会高兴，或会悲哀，有什么愿望，相信什么，等等。但是，很明显，我没有把它们的悲哀看得同人的悲哀那样深刻、复杂。

桑迪：但你不会认为这是“模拟的悲哀”，对不对？

帕特：当然不会。我认为这是真实的悲哀。

桑迪：我想我很难避免使用那些具有目的和心理色彩的词汇。我想使用这些词汇完全是正当的，当然，别用得太过分了。只是这些词汇用在对弈程序上同用在人身上没有同样丰富的含义。

克里斯：但我还是不明白，智能为什么非得要有感情掺杂在里面呢？难道说不能想象一种只会计算而没有感情的智能吗？

桑迪：你的问题有两种回答：第一种，任何智能都必须有动机。许多人以为机器的思维比人类更“客观”，这不是事实。机器观察客观事物时，它必须集注意力于某事物之上，并将

它纳入某些预先想好的范畴之中，这同人类的思维过程是差不多的。这就是说，它对事物有取舍，对事物有轻重之分。在每种信息处理阶段都是这样的情况。

帕特：你想说明什么问题呢？

桑迪：就拿我作例子吧。你可以说我正在提出某些智力性论点，这样做不需要任何感情。但是什么原因使我关心这些论点呢？我为什么如此强调“关心”呢？因为我在我们的讨论中动了感情！人们出于信仰，并非出自空洞、机械的反射而相互交谈。即使是高度理智型的对话也有极深刻的感情背景。人们的对话都受情感潜流的影响。因为你可以看到，你说话时希望有人在听，有人在理解你，会对你的言谈表示敬佩。

帕特：你的话听起来好像是，人们在说话时就需要别人感兴趣，不然话就没法谈。

桑迪：对！要是我没有兴趣，我干吗要去跟别人说话？而兴趣就是潜意识倾向性的总称。我说话时，我所有的个人倾向都在起作用，而从表面上看，你觉得那是我的风格和我的个性。这种风格完全是从许许多多的倾向、偏见和学习中产生出来的。要是有一百万个这样的倾向相互结合，你就会得到一个相当于许多欲望的东西。这完全是积累而成的；从这一点我又想起了另一个问题，那就是没有感觉的计算。当然，这是存在的，在点钞机里，在袖珍计算器里，这种情况是存在的。我还认为今天所有的电脑程序也是这样。但是，最终当你将足够的没有感觉的计算集中在一个巨大的协调组织里的时候，那么，你就发现了有另一个层次特性的东西了。你会看到，事实上你不得不看到，此时此刻这已不再是一堆没有感觉的计算了，而是一个由倾向、欲

望和信念组成的系统了。当事物变得很复杂时，你不得不改变描述的层次。在某种程度上说，这种情况已经发生了，因此，我在描述对弈程序和其他机械思维的试验时，就用了像“想要”、“思想”、“试图”和“希望”这样的词汇。丹尼特将观察者的这种转换层次称为“采取有意的姿态”。照我的看法，只有当程序自己对自己采取有意的姿态时，人工智能中真正有意思的事情才会发生。

克里斯：那是一个非常奇特的跨层次的反馈环路。

桑迪：肯定是的。当然，我认为现在就对计算机的程序采取这种有意的姿态还为时过早。至少我是这样认为的。

克里斯：我觉得有一个相关的问题十分重要。除了对人以外，你还对别的什么生物合理地采取这种态度？

帕特：我认为对哺乳动物也可以这样。

桑迪：我同意。

克里斯：这很有意思！桑迪，这怎么会呢？你当然不会认为猫或狗也能通过图林测试。但你不是认为图林测试是试验思维存在唯一的途径吗？你怎么能同时持这些观点呢？

桑迪：嗯……是的。我必须承认图林测试对象在一定层次上有意识。当然，有些有思维的生物也会通不过这个测试，但反过来说，只要能通过测试，依我之见，它就是真正有意识和有思维的生物。

帕特：你怎么能说一台计算机是有意识的生物呢？如果我说的话听起来像陈词滥调，我应表示抱歉。但当我想起有意识的生物时，怎么也不会将它们同机器联在一起。在我看来，意识总是和温软的肉体联在一起的，尽管有时有点傻，但还是健全的。

克里斯：一个学生物的人说这样的话，的确让人觉得有点别扭。

你是不是完全用化学和物理的语言来探讨生命，让一切神秘的东西都荡然无存。

帕特：不完全是。有时候，化学和物理反而让人觉得有些东西神秘莫测！不管怎么样，我不会总是把我的科学知识和内心情感联系起来的。

克里斯：我想我也是这样的。

帕特：那么，你对我这种固执的偏见怎么看呢？

桑迪：我想透过你"机器"观的表面，找到里面直观的意识，这种意识虽然不易被人发现，但却左右看待你的观点，我想我们都从工业革命的时代起继承了这样一个陈旧的观念，认为机器无非是在"隆隆"作响的引擎的驱动下，笨头笨脑地进行工作的一堆钢铁。也许那便是计算机的发明者查尔斯·贝比奇怎么会将人也看作是一堆钢铁的原因！毕竟，他将自己伟大的多齿轮计算机叫作分析引擎。

帕特：当然，我看人不是一台蒸汽铲土机，也不是一个电动开罐头的工具。人的身上有一种东西，一种非同寻常的东西，嗯，对了，在他们身上有一团火，一种充满活力的东西，那团火是不可预测的，飘忽不定的，但它是充满创造力的东西！

桑迪：你说得太好了！我就想听到你这种话。这才是有人味道的想法。你说人身上有一团火，这使人联想起蜡烛、火焰和天空出现闪电的雷雨。但你知道吗，在计算机的控制台上可以看到这种壮观的景象？一道道闪电构成了一幅壮观的景象。这同没有生命的金属堆相差十万八千里！上帝，那的确像火！"机器"这词为什么不让你联想起闪电的景象而是联想起蒸汽铲呢？

克里斯：桑迪，这是个美好的意象。这使我改变了对机械的看

法，机械不再是以物质为主，而是以模式为主。这使人试图把自己脑子里的思想——甚至还包括我现在的思想——想象成无数在跳跃的小脉冲。

桑迪：你将自己的思想想象成跳跃的脉冲，倒是挺有诗意的自我描绘。

克里斯：谢谢。但是，要叫我相信机器就是我，我还无法完全接受。我承认我对机器的看法颇有点不合时宜和下意识的味道，但我觉得，这种根深蒂固的看法很难转眼间就改过来。

桑迪：不过，我觉得你这个人还是心胸豁达的。说句老实话，我一方面十分同情你和帕特对机器的看法，但另一方面又不愿把自己当作机器。要是把你或我这样有感觉的生物看成完全是由电工线路中产生出来的，这是一个奇怪的想法。我说这话，让你感到意外吗？

克里斯：的确使我感到意外。你实话直说，你相信计算机有智能还是没有？

桑迪：这看你是什么意思了。我们都听到过"计算机能否思维？"这样的问题了。这个问题本身会有好几种解释（当然，光是"思维"一个词就有许多种说法）。这要看你对"能"和"计算机"这两个词怎么理解。

帕特：你又回到文字游戏上去了……

桑迪：没错，首先，这个问题可能是这个意思，"当代的计算机现在是否能够思维？"对这个问题我会理直气壮地回答，"不能"。其次，这个问题也可以理解为："如果程序设计得当的话，当代的计算机现在是否具有潜在的思维能力？"这种理解具有一定的合理性，但我的回答是，"大概不能"。真正的困难在于你对"计算机"这个词是怎样理解的。照我

的看法，“计算机”使人想起我刚才说过的那个形象：没有生气的装有空调的房间里，冷冰冰地躺着一些矩形的金属箱子。但恐怕随着计算机日益普及，其结构和性能的不断改进，这个形象最终会过时的。

帕特：但是，你认为计算机的改进需要一个很长的过程吗？

桑迪：当然需要，改进计算机决不是一夜之间所能完成的工作。当然，先进的计算机也是会出现的，其性能也许会大不一样。这样先进的计算机也许就不再叫计算机了。如同生物一样，计算机进化的系谱中也会有许多分枝。会有商业计算机、学童计算机、用于科学计算的计算机、用于系统研究的计算机、模拟计算机、火箭发射计算机等等。另外，还会有智能研究计算机。我感兴趣的是智能研究计算机，因为这类计算机灵活性最大，人们绞尽脑汁想让它们智力超人。我觉得它们完全有可能超出传统的计算机形象。也许在不久的将来，它们会配备最基本的感觉系统，首先，主要是具备视觉和听觉系统。它们需要有移动和探索的能力。它们在形体上也必须十分灵巧。总之，它们必须越来越像动物，具有越来越多的自主能力。

克里斯：这有点像电影《星球大战》中的 R2D2 和 C3P0 型机器人。

桑迪：当我想象智能机器时，我可没这么想过。因为这些机器人太愚蠢了，完全是剧作家们的凭空想象。连我自己也没有想好，这类计算机究竟会是一副什么样子。但是，如果你要想象一下的话，你千万不要受习惯的束缚，把这些具有人工智能的东西想象成笨手笨脚的蠢家伙。一切机器所共同具有的唯一特点就是最根本的机械性。这样说也许会使你觉得太冷漠，太死板了，但说到机械性，又有什么能

比我们细胞里的DNA、蛋白质和细胞器的活动更富有机械性呢？

帕特：对我来说，细胞里的活动是“湿的”、“滑的”，而机器里的活动却是又干又硬的。这是因为计算机一般不会犯错误，它们总是遵照人们的旨意行事。或者说这至少是我个人对计算机的看法。

桑迪：你的话很有意思。刚才你还在说像“一团火”，现在你又改口说什么“又湿又滑”之类的话。这不是表明我们有时会陷入多么极端的矛盾？

帕特：我不需要你的讽刺。

桑迪：我绝没有讽刺你的意思。我**真的**认为这很有意思。

帕特：这只不过是人类心灵的易滑性质的一个例子——这里说的是我的心灵。

桑迪：没错。但你对计算机的看法有点墨守陈规了。计算机当然也会犯错误，我指的不是硬件。就拿现代进行气象预报的计算机来说吧，它也会作出错误的预测，尽管它的程序毫无差错。

帕特：但这只是因为你输入了错误的数据。

桑迪：并不是这个原因，而是因为气象预报实在太复杂了。程序所使用的数据有限——当然这些数据完全正确——从这些数据中推断出气象的变化情况，有时就会作出错误的预报。这同田里干活的农民看着天上的云说，“我看今天晚上会有小雪”无多少区别。我们先在头脑里形成事物的模型，然后用这个模型来猜测世界会怎么变化。我们必须用头脑里的模型进行判断，也许这种模型并不是客观事物真实的反映。如果这种反映太不准确，那么，进化就会把我们淘汰——说不定会因为花了眼一脚从悬崖上摔下去。计算机也

同样。唯一不同的一点在于，计算机的进化速度会很快，因为计算机科学家们千方百计想尽快制造出人工智能；而这要是靠自然的话，则是偶一为之的。

帕特：那么，你的意思是说，计算机越先进，犯的错误就会越少啰？

桑迪：实际上正好相反。计算机越先进，它们就越能处理实际生活中许多乱七八糟的事情，这样一来它们也就越容易获得不准确的模型。所以对我来说，犯错误完全是高智能的表现。

帕特：我的天哪——你真让我无所适从了！

桑迪：我想我大概是一个人工智能特殊的拥护者。从某种程度上说，我有点脚踏两只船。机器再聪明也不会跟人一样，除非它们身上出现了生物般的湿润性和圆滑性。当然，我不是说机器内真的出现了水份，我的意思是说软件有灵活性。但是，不管像不像生物，智能机器毕竟还是机器。人类会设计它们，并把它们制造出来，还得发展它们！我们知道机器是怎么工作的——至少在某种意义上。可能没有一个人能真正了解机器，但是人类作为整体是能知道机器是怎样工作的。

帕特：你原来想两者兼得呀！

桑迪：你也许说得对。我想说明这样一个问题，人工智能不仅仅具有机械性，同时也具有有机性。生物机制里令人惊奇的灵活性也会在人工智能中出现。机制就是机械装置。DNA 和酶这样的东西实际上也是机械的、刚性的、可靠的。帕特，你不认为是这样吗？

帕特：没错，是这样。但是，当它们在一起工作时，许多意料不到的事情发生了。许许多多复杂事物和形形色色的行为

模式结合起来便成为一种灵活易变的东西。

桑迪：在我看来，从分子的机械阶段到细胞的生命阶段的转变简直是难以想象的。也正是由于这一点，我相信人就是机器。这种想法一方面让我觉得有点不舒服，但同时又让人无比振奋。

克里斯：如果人是机器的话，那为什么说服人们承认这个事实却又如此困难呢？很明显，如果我们真是机器，就应该发现自己身上的一些机械特征。

桑迪：这里有些感情因素。要是有人对你说，你是机器，这就是说你无非只有身体，这一下子使你想到了自己会死。这是人所难以接受的。但是，要是抛开感情不谈，认为自己就是机器，那么，你就得从最基本的机械阶段一下子跳到复杂的生命阶段。如果中间还有许多过渡阶段，它们就像一面屏障，使机械的特征几乎看不到了。我想，智能机器在我们眼里就是这样一种情况，在它们自己眼里也是如此。

帕特：我以前听说过一个笑话，说的是一旦智能机器出现以后会出现的情形。到了那个时候，你把智能输入你要控制的机器以后，它们的行为就无法预料了。

桑迪：也许它们身上会出现一小把奇特的"火"？

帕特：有可能。

克里斯：这个笑话有什么好笑的？

帕特：请想一想军用导弹。导弹中的目标跟踪计算机越复杂，它们的活动就会越难预料。最后，导弹会自认是和平主义者，会掉头回家落地而不爆炸。这样的话，将来战争中会出现一些"聪明的子弹"，它们中途掉头回家，因为它们不愿自杀！

桑迪：你说得真妙。

克里斯：我可不这么乐观。桑迪，究竟什么时候能发明出智能机器？我想听听你的看法。

桑迪：也许不会太久，也许在不久的将来，我们就会看到类似人类智慧的人工智能。但要想复制人类的大脑还不太现实，因为这实在是太复杂了。这仅仅是我个人的看法。

帕特：你认为计算机程序能顺利通过图林测试吗？

桑迪：这个问题很难回答。我看，说到这个问题，通过测试是有各种不同的程度的。这并不是黑白分明的事情。首先，这取决于提问者是谁。要是他是个大傻瓜，那他就会轻而易举地被现代的程序蒙骗过去。但是，还有一点也很重要，这取决于你能探究到多么深。

帕特：这样，图林测试就可以分成几类。有1分钟的，有5分钟的，有长达1小时的，等等。要是有个官方机构定期举行竞赛，就像年度计算机对弈大奖赛那样，看看哪个程序能通过图林测试，那不是很有意思吗？

克里斯：哪个程序能同一些著名的裁判员对答如流的时间最长，那它就将是冠军。如果哪个程序能蒙骗一个著名的裁判员长达10分钟之久的话，那它应获得一个大奖。

帕特：程序得了大奖有什么用呢！

克里斯：噢，帕特。要是程序能聪明地蒙骗裁判员，难道它还不会享受大奖吗？

帕特：当然会，要是这个大奖就是同所有的提问者在晚会上翩翩起舞的话；它尤其会享受这种奖励！

桑迪：我当然愿意看到举行这样的竞赛。要是看到那些名列榜首的程序的感人舞姿，那一定会很有意思！

帕特：你很悲观，对不对？要是提问者很老练，你认为计算机程序能顺利通过一个长达5分钟的图林测试吗？

桑迪：我对此抱有很大的怀疑。首先，迄今为止还没有人真正试验过。但是，现在有一个叫作“帕里”的程序，它的设计者声称，它已经顺利通过了一种初级的图林测试。在一系列遥控的提问中，帕里成功地蒙骗了数位精神病学家，他们事先知道他们的对话者有可能是计算机，也有可能是患狂想症的精神病人。在这之前还有另外一个更简单的测试方法，依照这个方法，精神病学家只能看到一些简短对话的记录，然后决定哪些对话人是真正的狂想症患者，哪些是计算机。

帕特：你的意思是说，精神病学家没有提问题的机会？这里就有一个严重的不足，而且与图林测试的精神不符。要是你让某人猜我的性别，我什么都不对他说，只让他看几段我写的话，他会怎么判断呢？当然非常不容易！所以，后来的那个方法比早先那个有进步，我很高兴。

克里斯：你通过什么办法使计算机像一个狂想症患者？

桑迪：我没说计算机真的像一个狂想症患者，我只是说某些精神病学家在不寻常的环境中，以为这是精神病人所为。在这种假图林测试方面有一件事使我很担心，这就是帕里的行为方式。为什么说“他”——他们这样称呼——像精神病人是因为他的防卫意识很强，在谈话中对不感兴趣的问题趁早避开。最重要的是，他很能自控，因此没有人能够真正地试探“他”。因而模拟一个狂想症患者要比模拟一个普通人容易得多。

帕特：别骗人了！这使我想起计算机程序最容易模拟的那种人的笑话来。

克里斯：什么人？

帕特：紧张症患者，这种人整天闲坐着。什么事也不干。连我

都能编一个程序来模拟这种人！

桑迪：帕里还有一个有趣的特征，他自己不会造句，但在它里面存储了大量的句子，它只是选择其中最合适的一句，作为对输入句的反应。

帕特：这太奇特了！但要大规模这样做便不行了，对不对？

桑迪：是这样。在一场谈话中为了正常应付所有可能听到的句子，你要存储的句子将是个巨大的天文数字，这实际上是不可想象的。而且为了在存储中找这些句子，所需要的索引也肯定极为复杂……要是谁认为程序在存储里找所需要的句子就像自动电唱机找一张唱片那样容易的话，那么，这个人肯定没有认真研究过图林测试。这里有一点很有意思，当图林测试的反对者们高唱反调时，他们就常常援引这种不可实现的程序作为反证。他们要你想象的不是一台真的智能机器，而是一个庞大笨重的机器人，用单调的音调吟诵着已存储的句子。他们还认为，尽管机器已经在进行相当灵活的智能活动，它的机械性特征还是能够一目了然的。然后，他们就会说："你瞧瞧！这不就是台机器吗？不过是个机械的玩艺儿，根本没有智能！"而我对这个问题的看法恰恰相反。要是一台机器的能力可以和我一样顺利地通过图林测试，那么，我丝毫也不会觉得这是对人类的嘲弄和威胁，相反，我会和哲学家雷蒙德·斯马林一起拍手唱道："机器啊，你是多么美妙！"

克里斯：如果在图林测试中你只向计算机提一个问题，你问什么呢？

桑迪：嗯……

帕特：就问："如果在图林测试中你只向计算机提一个问题，你问什么呢？"行吗？

D·C·丹尼特

反　思

图林测试规定，参加模拟游戏的被问者必须同提问者隔离，他们之间只能通过书面语言进行交流。对此许多人都持有保留意见。因为，这样做是不是故意方便了被问者向提问者隐瞒事实？如果把被问者关在“黑箱”中，提问者除了他们有限的“外部行动”之外，别的一无所知（在这里，就是指打字机打出的回答），那么，图林测试只能教条地按照某种行为主义的方式进行，或者（更糟的是）按照操作主义的方式进行，或者（还要糟的是）按照确证主义的方式进行。（这三种主义不久前还是可怕的怪物，现在已经被科学哲学家完全摈弃了——但是，那里是什么讨厌的声音？它们会不会还在坟墓里动弹？我们早该拿刀子直插它们的胸膛了！）图林测试是不是约翰·塞尔说的“操作主义的花招”之一例？

很明显，图林测试强调心灵的本质这个问题。图林认为，被试者的脑袋里生着什么样的灰白质（如果有的话），是什么个样子，什么气味，都不要紧；重要的是它是否有能力进行有智力的活动，或者说表现这种行为也可以。图林提出的那个游戏即模拟游戏决非神圣的，只不过是谨慎地选择的一般智能测试。他的假设是：如果被试没有能力作出一系列有智力的反应，那么，他就不可能赢得模拟游戏的胜利，也就通不过图林测试。幸好图林没有把挫败世界象棋冠军作为其测试智力的标准，要不然的话，他必然会遭到理由充足的强烈反对。现在看来，要想制

造一台只能挫败世界象棋冠军，但除此以外别的什么都不能做的机器，还是非常可能的。如果图林把智偷大不列颠王冠或是和平解决阿拉伯与以色列争端作为衡量智力的标准，那就不会有什么人反对说，智力被图林"还原为"行动了，或者说他根据行为来下"操作式的定义"(当然，会有某位哲学家绞尽脑汁去设想一个完全不可能实现的方案，根据这个方案，王冠差不多是自己长腿跑到一个傻瓜的手心里的。这样，他就可以说，这个蠢才已顺利"通过"测试，因此，这就"证明"图林测试并不能有效地测验智能。当然，真正的操作主义者也会承认这个幸运的傻瓜是有智力的，因为不管他有多傻，他还是顺利地通过预定的测试。这也正是你不易找到一个真正的操作主义者的原因所在)。

为什么说图林测试所选择的方法比偷英国王冠或和平解决阿以争端要好呢？这是因为后面两种方式是不可重复的(一旦办成功，就不会再出现第二次了!)，难度也太大了(许多一副聪明相的人对此也无能为力)，同时也很难得出客观的结论。图林测试就像巧妙的打赌一样鼓励尝试。它看上去很公正，全凭可能性，最后的结论也很客观。但实际情况并非如此。事实上，图林这样说："我们为什么要喋喋不休地去讨论心灵或智能最终是什么这样的问题呢？实际上很简单，我们可以认为凡是能通过这个测试的任何东西都可被视为肯定是有智能的。然后，我们就可以考虑怎样去设计能顺利通过测试的机器这个问题。"但是，具有讽刺意义的是，图林并没有解决这个问题，他只不过设法转移了问题的方向。

图林测试是不是因为它的"黑箱"观点而容易受到批评呢？首先，正如霍夫施塔特在上面那段对话中指出的那样，我们都把别人看作黑箱，根据观察他人明显的智力行为来建立对他人

心灵的信念。其次，黑箱观点是一切科学研究的观点。在研究DNA分子时，我们从各个方面对它们进行探究，观察它们的反应。在研究癌、地震以及通货膨胀时，我们也采用了同样的方法。在研究宏观物质时，“窥视”黑箱的办法还是十分有用的。我们可以(例如，用解剖刀)把物体切开，然后就能观察物体内部的情况。我再就黑箱问题赘言一句。正如霍夫施塔特说过的那样，我们肯定会遇到这样一个问题：使用什么样的探测手段，才能同我们所要回答的问题密切相关？如果我们想解答的问题是某个实体是否有智力，那么，除了问一些关于日常生活的问题之外，我们想不出更直接、更能说明问题的探测手段了。图林的“行为主义”无非就是将近乎陈腐之言的结论同具有实验风格并且很方便的测试结合起来。

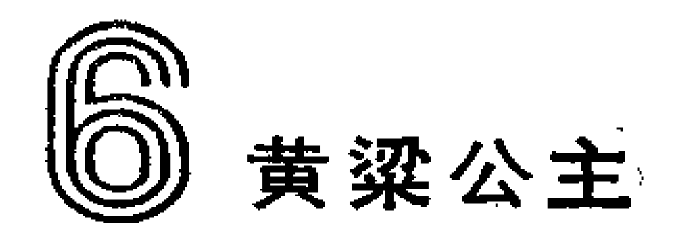
在这段对话中，霍夫施塔特还提出了另外一个问题，但并没有给予解答。问题涉及到计算机对客观事物的模拟。计算机模拟是对事物详尽、“自动”、多维的模拟，当然，模拟与现实之间存在巨大的差别，对不对？约翰·塞尔曾经这样说过：“谁也不会相信，我们能够仅凭计算机模拟乳汁分泌和光合作用的过程产生出牛奶和白糖来……”如果我们设计在数字计算机上模拟一头奶牛的程序，这头模拟的奶牛绝对不可能被挤出真奶来，最多只是模拟的奶。不管这奶模拟得有多像，也不管你有多渴，模拟的奶是不能喝的。

但是，假设我们用计算机模拟一个数学家，并且模拟得很逼真。我们会不会抱怨，我们所希望得到的是证据，可是天哪，到头来我们得到的只是证据的表现而已？但是，证据的表现也是证据，对不对？这取决于表现的逼真程度。漫画家在给科学家画漫画时，常在他们的身边画上一块黑板，再在黑板上涂上一些公式或数字。尽管这些公式或数字在普通人的眼里似乎很

“真实”，实际上是漫画家信手画来，毫无意义。如果模拟的数学家只能给予一些像漫画中那样的假证明，它还是具有一定的理论意义。因为它模拟了数学家们的侃侃而谈或心不在焉。另一方面，如果模拟的目的是为了模拟一位优秀数学家所作的证明，那么，它就会是一位同数学家一样有价值的证明提供部门的“同事”。似乎这便是抽象（形式产品如证明或歌曲）与具体（如牛奶这样的物质产品）之间的差别（请见下一章“黄粱公主”）。心灵应属于哪一类呢？心灵像牛奶还是像歌曲？

如果我们认为心灵的产品像身体的控制之类的东西，那么，它的产品似乎十分抽象。如果我们认为心灵的产品是一种特殊的物质或数种特殊的物质，比如，许多许多的爱、一丁点儿痛苦、一些心醉神迷、一些所有棒球手都充分具备的欲望，那么，它的产品就十分具体。

在针对这个问题展开辩论之前，我们应该先慎重考虑一下：如果我们遇见了对任何具体物质或现象的真正详尽而出色的模拟，构成这种区分的原则在我们必须推广的界限上是否很清楚。任何实际连续的模拟都将在这样或那样的硬件中得到具体的“实现”，因此，模拟的工具本身也会对现实产生某种影响。假如模拟的事件与真实的事件对现实能产生同样的影响，那么，如果你还坚持说模拟毕竟是模拟，也未免太固执了。在下一章里，这个问题被说得妙趣横生，这也是全书反复讨论的一个主要问题。



6 黄梁公主

斯坦尼斯瓦夫·莱姆

“有件事……什么事忘了，”国王回到做梦大臣面前说道，“苏布蒂利恩，可你为什么要这样用一条腿跳，另一条腿抱着?”

“没——没什么，殿下……有点关节炎……一定是天变了。”狡猾的特豪马图尔盖结结巴巴地说，接着，他又继续引诱国王再试一个梦。国王日佩鲁普斯沉吟片刻，查了查目录，选中了“黄粱公主的新婚之夜”。他梦见自己坐在火炉旁，手捧一卷奇怪的古书，书是用深红的墨水在烫金的羊皮纸上写成的。书中用优美的语言描绘了一位黄粱公主，她在五百年以前统治着蒲公英王国；书中还描写了她的冰凌森林、螺旋宝塔、鸣叫的鸟舍以及百眼宝库，但是，主要讲的是公主的美貌和品德。日佩鲁普斯渴望亲眼看见这个可爱的人儿，他的胸中燃起了欲望之火，心灵在燃烧，闪闪发亮。他出门登程，在梦中的每一个角落搜寻黄粱公主，但哪里也找不到她；确实，只有最古老的机器人才听说过这位公主。国王因长途旅行疲惫不堪，最后来到

了一座大沙漠，只见一个个沙丘都是由金子堆成的，国王发现了一座小屋。当他走近小屋时，他看见一位年迈的长者，身穿一件雪白的袍子。长者站起身来，说道：

“你在寻觅黄粱，可怜的人儿！可你明知道这五百年来她根本没有活过，因此你的热情都是徒劳的！我没法让你看见她的血肉之躯，只能让你看见她的一个数据摹本，她是一个数字模型，没有血肉，没有弹性，也不具有立体感，更不带有肉欲。我空闲时便一块块拼起来，成了一个黑盒，都在这里面！”

“啊，让我看她，请让我现在就看到她！”日佩鲁普斯颤抖着大声祈求道。长者略一点头，便在一册古书中寻找公主的坐标，将她和整个中世纪打在穿孔卡片上，写毕程序，通上电流，掀开黑盒的顶盖，说道：

“看！”

国王探身向前，举目细看，面前丝毫不差地呈现出模拟的中世纪图景，都是非线性的，两进制的、数据式的。那里有蒲公英王国、冰凌森林、有螺旋宝塔的宫殿、鸣叫的鸟舍，还有百眼宝库，黄粱公主本人悠闲地在模拟的花园里散步，在她俯身去摘模拟的雏菊，嘴里低吟着一首模拟的歌曲时，她身上的线路发出红光和金光。日佩鲁普斯再也控制不住自己的激情，跳上黑盒，不顾一切地想爬进那个计算机的世界。但是，那位长者立刻切断了电流，将国王推倒在地上。

“你这疯子！为什么尽想那些不可能的事情！物质的东西永远也不能进入那个虚无的系统，因为在这个系统里只有文字、不连贯的整数以及抽象的数字！”

“可我一定要，一定要！！”日佩鲁普斯神志不清地乱叫乱嚷，用头使劲撞着黑盒，直到金属外壳被撞得凹陷了。长者接着又说：

"如果你执意要这样做的话，我倒是有一个办法成全你，但是，你必须离开你现在的肉体，因为我得确定你身体各部位的坐标，然后一个原子一个原子地把你编成程序，再将模拟的你放在中世纪的世界里，这只是一个信息和表现的世界，你将留在那里面，只要电子通过这些线路从阳极移到阴极，你就可以一直在里面。只是，现在站在我面前的这个肉体的你将被消灭，这样，你将以特定的场和势的形式存在，一切由统计的、直观推断的和数字的东西构成！"

"这简直令人难以置信，"日佩鲁普斯说。"我怎么能肯定你的那个模拟的人是我而不是别人呢？"

"很好，我们可以试一试。"长者答道。他量了国王身上所有的尺寸，就像裁缝做衣服那样，只是比裁缝量得精确多了，因为每一个原子都经过了仔细的划分和测量。然后，长者将程序输入黑盒，说道：

"瞧！"

国王偷眼一瞧，看到自己正在黑盒里偎火读一卷古书，书中描写了一位黄梁公主，然后他夺门而去，四处寻觅公主，最后在金沙漠中找到了一座简陋的小房子，那里有一位全身雪白的长者，长者对他说道："你在寻觅黄粱公主，可怜的人儿！"等等。

"现在你信了吧？"长者说道，闭上开关。"这次我将程序编成你生活在中世纪，依偎在可爱的黄粱公主的身旁，与她一道共享不尽的模拟、非线性和二进制的美梦。"

"是的，是的，我明白，"国王说道。"但是，还有个问题，在黑盒里的只是像我，而不是我，因为我在这里而不是在黑盒里！"

"可你不会久在此地，"长者答道，满脸微笑，"因为我就要……"

长者从床下抽出一把很沉但又很管用的铁锤。

"当你被心爱的人紧紧拥抱，"长者告诉国王，"我将考虑到你只有一个，不能既在这儿，又在那儿，在黑盒里——我将用一种原始、古老但又十分灵验的办法，是不是可以稍稍弯下身子……"

"先让我再看一眼你的黄粱公主，"国王说："弄确实……"

长者打开黑盒的顶盖，又显出了黄粱公主的倩影。国王看着、看着，最后开口道：

"这本古书中关于公主的描写太夸张了。当然，她不错，可总比不上历史中记载的那样美丽。那好，再见，尊敬的长者……"

国王扭头就走。

"你去何方？疯子！"长者紧握锤把，大声喊道，因为国王已走出门槛。

"哪儿都行，除了这个黑盒，"日佩鲁普斯国王答道，匆匆离去。就在此时，国王从这场美梦中骤然惊醒，他发现自己站在门厅前，面对着万分失望的苏布蒂利恩。后者因为就差一点没把国王锁在黑盒内而感到失望。大臣特豪马图尔盖本来可以将国王永远禁锢在那里……

D·R·霍夫施塔特

反　　思

在这本书中，我们选了波兰作家和哲学家斯坦尼斯瓦夫·莱姆的文章，这是第一篇。我们用了迈克尔·坎德尔的译文。在

评论莱姆的观点之前，首先必须感谢坎德尔天才的翻译，他将波兰语中妙趣横生的巧妙对话译成了同样有趣的英语。在《希伯里亚德》的全书中(上面的故事就选自此书)，坎德尔的高超的翻译水平始终如一。读着这样精美的译著，我们不禁感叹当前用机器翻译程序取代人工翻译还为时太早。

对于本书中提出的那些问题，莱姆毕生都感兴趣。莱姆通过直觉和文学的手段来说服读者相信自己的观点，其效果比起刻板的科学论文和神秘推理的哲学专著要好得多。

至于他的故事，我们读一读便明白了。我们只想知道一件事情：一首模拟的歌与一首实在的歌之间究竟有什么区别？

7 猩猩玛塔的灵魂

特雷尔·米达纳

贾森·亨特致了谢，长长地松了一口气，然后传唤下一个证人。

亚历山大·别林斯基博士是动物心理学教授，他长得矮矮胖胖，举止唐突而又一本正经。一上来，他先向法庭出示了自己获得优秀学术成果的证书，以证明他是这方面的专家证人。完了以后，亨特要求法庭允许进行较复杂的证明。

法庭对是否应允许这样做小议了片刻。由于莫里森没有反对意见，所以，尽管费曼表示保留意见，法庭还是同意了。不一会儿，法警带进来两名研究生助手，还推进来一辆装着各种电子仪器的小车。

按照惯例，法庭只限于书面记录，那种利用电子仪器的证明法只是这几年才允许的。近年来，为了加快法庭程序的进行速度，颁布了一些专门的法律，允许书记员用录像机录制这样的证明，以作为正式记录。但是，当费曼看到一个助手忙着安装

电子仪器，另一个离开一会儿以后又领进来一只猩猩时，他开始后悔现代化的到来。

猩猩看上去十分紧张，因人多而害怕起来，它紧紧地依偎着领它进来的那位研究生助手。它一看见别林斯基博士，就立即跳进了证人席，脸上明显地露出亲热的表情。在亨特的示意下，别林斯基将猩猩介绍给法庭。这是他最近实验用的一只猩猩，他为了研究总共用了二十只猩猩，其研究成果最近已写成著作出版了。随后，经亨特的要求，别林斯基叙述了这些实验的情况。

“许多年以来，人们一直以为动物由于没有发达的大脑，所以不具备像人类那样的语言器官。但是，在六十年代初期，一些动物心理学家提出，猩猩之所以不能说话的唯一原因是因为它们的声带不健全。为了证明这一理论，他们设计了一些简单的不涉及话语的符号语言。他们试验了彩色卡片、画片、磁石板、键盘装置，甚至是国际符号语言，并且都取得了一些成功。

“这些实验证明了符号语言不仅仅为人类所有，同时也表明了即使是最聪明的动物，它们的语言功能也是十分有限的。一位十分有才华的大学生后来发明了一个计算机程序，这个程序能够模拟最聪明的猩猩在语言方面的任何成就，学术界对动物语言实验的兴趣也随之大减。

“然而，似乎这些动物可能受到了过去实验条件的限制，正和它们过去受到不发达声带的限制一样。人的大脑中有一个语言中心，这是一个特化的区域，专门用来理解和创造人类语言形式，猩猩是在自然状态中相互交流的，也有一个特化的大脑区域，含有使它们能吱吱叫和大声嚎叫的自然系统。

“我发现，先前的语言实验只重视猩猩的手势而不重视它

们的声带，也忽视了它们的自然语言中心。于是，我打算研究这个自然语言中心，但仍然不注重它们不发达的声带，结果我取得了成功，所用的试验仪器现在就放在诸位面前。

“诸位请仔细看看玛塔头部的左面，这里有一个圆形的塑料帽。在塑料帽下面，我将一个电子连接器永久性地嵌在它的脑壳里。连接器上有一些电极通向大脑里面。我们的电子仪器与玛塔的脑袋相连接，这样就能控制它的语言中心的神经活动，并能将神经活动翻成英文。

“玛塔是一个只装了7个电极的猩猩，所以它的反应很慢。它通过刺激装在头部的电极而‘说话’，但它自己并不知道这一点。电极发出的信号由一台微型计算机破译以后，把它选定的词语输出给声音合成器。这种技术使玛塔产生一种自然反馈的功能。当我们接通它的晶体管声带时，它就会发出人一样的声音，只是语法差些和没有词形变化。

“但是，我已经说过，玛塔并不是一个出类拔萃的学生，所以对它不要抱太多希望。尽管它脑袋中的7个电极系统能够破译出128个单词，可实际上它只学会了53个。别的猩猩要比它好多了。学得最好的是一个装有9个电极的雄猩猩，它的系统能破译512个单词，它掌握的单词达407个。尽管如此，我相信你们会发现玛塔很健谈。”别林斯基最后补充道，伸手去摸它的连接电缆。

当别林斯基博士开始为玛塔接通连接人类语言世界的线路时，它显得十分高兴和激动。它上蹦下跳吱吱地叫着，别林斯基接过他助手递过来的线路，然后坐下，打开玛塔头上的保护帽，将连接器的两端合上。一合上电闸，玛塔就蹦了起来，看上去，它并没有注意到连着它脑袋的线路，它伸手指向博士手中捧着的一个小盒。

“对玛塔来说，”他解释道，“语言几乎是一种无休止的活动，因为它的电子声带不会感到疲劳。为了输入一个词，我用这个控制装置其实是先让它停止说话。”

“好了，玛塔开始吧，”博士说着，打开了它的声音开关。

装在仪器车上的一个小麦克风立即嚷嚷起来：“喂！喂！我玛塔玛塔，快乐的猩猩。喂喂……”

就在法庭上人们大为震惊的时候，猩猩的说话声卡嗒一声切断了。刚才玛塔的嘴巴一张一闭，麦克风中传出它的模仿性感的女声的景象让人看了简直难以置信。

它的教师然后继续他的表演：“玛塔多大了？”

“三，三，玛塔，三……”

“很好。现在放松，玛塔，别紧张。我是谁呢？”他问道，手指着自己。

“别林斯基，人，好别林斯……”

“那么，那些人是谁？”他又问，用手向人头攒攒的法庭扫了一遍。

“人，人，人们，好人们……”

博士中断了猩猩的说话，转向辩护律师，示意自己已准备好了。

亨特站了起来，提出了第一个问题：“在你看来，这只野兽有智力吗？”

“从‘智力’这个词的广义上说，我认为它有智力。”

“从人类意义上说，它有智力吗？”亨特又问。

“我认为有的，但是，要是你认为它有智力，你就得将它当人看待，和它交谈，和它玩耍。为此我特地带来一盒它最喜欢的玩具。它会把它的有限注意力集中在我身上，或集中在任何一个拿了它宝贝玩具的人身上。我建议你自己试试看。”

莫里森瞟见费曼法官正注视着他，那意思是说叫他投反对票。于是，他就老老实实地说："我反对，尊敬的法官。至少亨特先生应该向我们证明这个证明与本案有关。"

"亨特先生，你有什么意见？"费曼问道。

"这个证明与本案有关，马上就会见分晓。"

"如果与本案无关的话，"费曼说，"请放心，我会让人将这段录像擦掉的。现在开始。"

亨特打开了玛塔的盒子。这是一个大珠宝盒，被涂成红色和银色，亨特看看里面的东西，抓起一支包着玻璃纸的雪茄。玛塔马上尖声尖气地喊了起来："雪茄，别林斯基，坏，坏雪茄……"接着又叽哩呱啦地叫了一通，还拧拧鼻子表示强调自己的话。

"玛塔，你的玩具盒里放一支旧雪茄做什么？"亨特问道。

"什么！什么？什……"它反问道，博士连忙中断了它。

"这个问题对它来说有点太难了。能不能将问题简化为主要的词和较短的动词？"博士这样建议。

亨特照办了："玛塔吃雪茄吗？"

这次玛塔有了反应："不吃，不吃雪茄。吃食物食物，抽雪茄。"

"很好，博士，"亨特称赞别林斯基。随后他又转向莫里森："原告是不是想检查一下证据？"

莫里森犹豫片刻以后同意了，他接过玛塔的玩具盒，非常不情愿地从里面拿出一只玩具熊让玛塔去认，它马上愤怒地蹦了起来，它那拟人的声音霎时也变得怒气冲冲：

"人，坏，坏，别拿熊，玛塔的熊，快来啊，别林斯基，快来啊，帮帮玛塔拿熊，快来啊……"

它的声音又被关闭了，恢复了平常的吱吱声。博士连忙向

大家解释玛塔的妄想症。“先生，它对您抱有一定的敌意。说老实话，我对您表示同情，我也明白，许多人和您一样，对动物能说显示智力的话感到非常不适。但它现在有点恼怒了。或许，要是能换一个人同它谈话，那会……”

“我想试试，”费曼法官打断了博士的话。众人当然表示同意，莫里森将盒子送到法官面前，玛塔此刻怒气已消，对于原告的不悦并不在意。

“玛塔饿吗？”费曼问道，因为他看见盒子中有几个熟香蕉和一些糖块。

“玛塔吃，现在，玛塔吃……”

“玛塔想吃什么呢？”

“玛塔吃，现在……”

“玛塔想吃糖吗？”

“糖，糖，想的，糖……”

他伸手拿出一个香蕉递给玛塔，它马上灵巧地接住，剥去皮，一口塞进嘴中。趁它在大吃大嚼之际，别林斯基又打开开关，正好听见一串“快乐的玛塔”的话，话语声使猩猩也吃了一惊。吃毕，它又面对法官，嘴巴一张一合，这时，博士又扭亮了开关：“好，香蕉，好，香蕉，谢谢你，人，糖，现在，糖，现在。”

费曼对此感到很高兴，他伸手从盒子里取出糖块来。玛塔接过糖，但并没有立刻吃，而是伸手指向别林斯基手中的电闸盒，表示要别人听它讲话。

“雪茄，雪茄，玛塔要雪茄……”

费曼找到了那支雪茄，递给猩猩。它接过来，放在鼻子下嗅了嗅，又递还给他。“好，好，人，吃，别林斯基，雪茄，谢谢你，谢谢你，人……”

费曼完全被猩猩的智力和天真烂漫吸引住了。玛塔也受宠若惊，表示亲热，这使法庭的气氛变得很活跃。但亨特不愿在这上面耽搁太久，过了几分钟以后，他打断了费曼与玛塔的对话：

“也许我该开始作证词了吧，尊敬的法官？”

“当然可以，”法官表示同意，非常不情愿地把同他一起坐在法官席上的玛塔抱了起来。

“别林斯基博士，”待玛塔坐定之后，亨特继续发言，“你可以就这个猩猩的智力简要地谈谈你的科学结论吗？”

“它的心灵是与我们不一样，”博士说。“但只是程度上的差异。我们的大脑体积更大，身体的适应能力更强，因此，我们很优越。但是，人和猩猩之间的差别也许会被证明是很小的，这会使我们很难堪。我认为，尽管玛塔有许多缺陷，但它仍具备与人类相同的智力。”

“你能否在人和猩猩的心理之间作一明确的区分？”

“我做不到。很明显，它比一般的人的智能要低。但比人类中的白痴要聪明，与大多数低能儿相类似。它另外有一个优点，比较讲卫生，能照顾自己和后代，而白痴和低能人做不到。我不愿意在它和人的智力之间作出明确的区分。”

亨特没有马上提出下一个问题。当然，他事先与博士一起共同准备了这个实验。为了完成今天的证明，他还要做一个实验，而这个实验本来是不大可能进行的。但是，他不知道博士是不是打算按原计划进行这个实验。事实上连他自己也闹不清楚究竟该不该做这个实验。然而，还有一件事：

“别林斯基博士，猩猩有类似人的智力，它是否能得到相应的像人那样的待遇？”

“不能。当然，我们对待实验用动物都很不错，但是它们的

价值只体现在实验潜力上面。就说玛塔吧，它已经活得够长了，它的实验使命早已完成，按计划它要被立即消灭，因为饲养它花的钱远远超过其实验价值。”

“可你用什么方法杀死这样一只动物呢？”亨特问道。

“现在有许多种无痛苦的快速方法。我想在玛塔最喜爱的食物里放入毒药，然后诱使它吃下，这种死法它自己预料不到。尽管这很残酷，但是，这样做它就不会预感到自己的不幸的命运。谁也逃脱不了死亡的威胁，但对这些简单的动物来说，没有必要让它们感到对死亡的恐惧。”说着，别林斯基从上衣口袋里掏出一块糖。

“你能在法庭上进行你的实验吗？”亨特问。

博士把糖块递给猩猩，费曼这才明白过来发生了什么事情。他大声惊呼，命令停止这一实验，但是，太迟了。

别林斯基从来没有亲手杀死过自己实验室里的动物，他总是让助手去干这种事。他眼睁睁地看着玛塔毫不怀疑地把糖块放进嘴里，嚼了起来，这时他如梦初醒。他扭亮开关，只听见：“糖，糖，谢谢你，别林斯基，快活，快活的玛塔。”

这次声音自行停止了。玛塔的身体抽动了一下，然后瘫软下来，静静地躺在它主人的怀抱中，死了。

玛塔是死了，可它的大脑并没有马上死亡。它一动不动的身体内某个电路的最后的感觉放电触发了一阵神经脉冲，译解如下：“伤害玛塔，伤害玛塔。”

有两秒钟整个法庭鸦雀无声。接着，任意触发的神经放电虽然与这头动物尸体无关，却又对人类世界发出最后一个搏动信号：

“为什么，为什么，为什么，为什么……”

电开关轻声关上了，结束了法庭作证。

D·R·霍夫施塔特

反　思

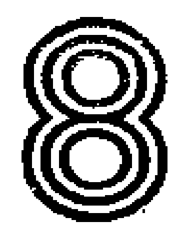
> 早晨在办公室处理公事。不一会儿，W·巴滕斯爵士把我们叫去，让我们看霍姆斯船长从几内亚带来的一头怪物。这是一只大狒狒，它在许多方面长得那样像人(尽管人们说确实有狒狒这种动物存在)，我真相信这是男人与雌狒狒生下的怪物。我相信它已经懂许多英语，我看可以教它说话、打手势。
>
> ——塞缪尔·丕普斯的日记
>
> 1661年8月24日

猩猩临死时的悲恸的、不明不白的叫声激起我无限同情，我们对这个无辜、可爱的生物很容易产生同情心。这一幕的可信性又在哪里？猩猩语言在过去十年里一直是个有争议的研究领域。从表面上看，猩猩和其他灵长目一样，能够接受许多词汇，甚至能达到数百个，有时候它们竟能使用巧妙的复合词，尽管如此，这并不能证明它们能够掌握一整套语法，并运用语法将词组合成表达复杂意义的句子。看起来猩猩只会随意地将词叠加起来，而不是将词连成句法结构。问题症结就在这里吗？在有些人眼里，问题就出在这里，因为没有句法结构就无法表达清楚复杂的意思。乔姆斯基等人坚持认为，人的本质就是有天生的语言能力，一种“基本语法”，所有的语言在深层结构上都有这种语法。照此说法，猩猩和其他灵长目由于没有这种基本语法，因此就同我们有本质上的区别。

还有的人认为，灵长目表面看来也使用语言，可是这种行为与人类使用的语言的行为大不一样。人类使用语言是为了交流，即将个人的思想变成共同的符号系统，而灵长目使用的符号对它们自己说来并没有意义，但是它们运用这些符号能达到希望达到的目的。从一个严格的行为主义者的角度看，像“意义”这样的心理因素都是假设的，如果依照这些假设的东西来区别外部行为，那是荒谬的。但是，这些实验以前确实做过，当然被试不是灵长目，而是中学生。参加实验的学生每人都分到各种形状的彩色塑料板，并要求他们将塑料板拼成一定的形状，能干者有奖。实际上，达到规定形状的拼排次序完全可以译解成非常简单的英语，然而，大部分学生都说他们从来没有这样想过。他们说，他们发现有些形状能拼成，而另一些形状则不行，仅此而已。在他们看来，这好比一场游戏，只不过将没有意义的符号颠来倒去地摆弄。这个实验结果十分惊人，它使许多人相信，猩猩语言的说法纯粹是拟人说动物爱好者的主观想象。但是，这场争论远远没有结束。

然而，不管这篇文选有多少实在论成分，里面却蕴藏着极深刻的道德和哲学上的问题。有心灵（即有智力）与有灵魂（即情感）之间有什么区别？两者是否可以独立存在？谋杀猩猩玛塔的理由是因为它不如人那么“有价值”。在这句话的背后还隐含着这样的思想，即玛塔的“灵魂不如人”。然而，智力是否真的可以成为灵魂的衡量标准？反应迟钝和年迈体衰的人的灵魂是否不如常人？音乐评论家詹姆斯·亨尼克在论述肖邦练习曲作品25第11号时这样说道：“灵气不足的人，不管你的手指多么灵巧，不要演奏此曲。”多么令人不可思议的声明！尽管有人会说这是势利眼，高人一等的优越感表现，但它还是有一定的合理性。但是，谁又能提供衡量灵气的标准呢？

图林的实验难道不是这样的标准吗？我们是否能够借助语言来衡量灵魂。毫无疑问，玛塔灵魂的某些性质已经通过它的大声而清晰的话语显示出来了。玛塔之所以富有感染力，一部分是由于它的形象(实际上，我们怎么知道这一点?)；另一部分是由于我们将它认作与自己同类；还有一部分是由于它那有趣的天真措词。我们觉得有义务保护它，就像有义务保护一个婴儿或幼儿一样。



三号兽的心灵

特雷尔·米达纳

“阿纳托尔的态度十分明确，”亨特说。“他认为生物的生命只不过是复杂的机器的一种形式。”

她耸耸肩膀，但并非很冷淡：“我承认他给我留下了极深刻的印象，但是我不能接受他的那种哲学。”

“你好好想一想，”亨特建议道，“你很清楚，按新进化论的说法，动物身体完全是由机械过程形成的。每个细胞都是一架微型机器，它们共同组成更庞大、更复杂的机体。”

德克森摇了摇头：“但是动物和人的身体不仅仅是机器。生殖行为本身使它们各不相同。”

“怎么，”亨特问道，“从一架生物机器产生另一架生物机器就那么奇妙吗？一个雌性哺乳动物怀孕、生殖时用不着比一家自动化工厂生产机器多费脑筋。”

德克森双眼发亮。“难道你认为自动化工厂生产机器时会有什么感觉吗？”她反问道。

“金属会受到高度挤压，机器最终会疲劳。”

"我说的'感觉'不是这个意思。"

"我也不是这个意思，"亨特表示同意。"但是，我们始终很难知道谁或是什么会有感情。在我幼年时呆的农场里，有过一头下崽儿的母猪，它总是把自己下的猪崽儿都给踏死——我想是碰巧。然后它就吃猪崽儿的尸体。你认为它有母亲的感情吗？"

"我谈的不是猪！"

"谈人也可以。你是否估计过究竟有多少婴儿被溺死在马桶里？"

德克森张口结舌，说不出话来。

沉默几分钟后，亨特接着说道："克兰对机器却另有一套看法。在他看来机器是一种生命形式，他自己就可以用塑料和金属来制造。而且他很诚实，认为自己也不过是一台机器。"

"大机器生小机器，"德克森的话中带着冷嘲热讽，"接着你会说他是一个母亲了。"

"不，"亨特说，"他是个工程师。制造的机器尽管比人体粗糙，也比简单的生物繁殖高级，因为至少它是思想过程的产物。"

"律师真难对付，"她承认道，但仍惊讶不已。"可我并没有涉及机器！从感情上说，我们对动物和机器是加以区别对待的，这很难用逻辑来说明。我的意思是说，我可以毫不在乎地打碎一台机器，却不忍杀死一只动物。"

"这你以前试过吗？"

德克森想了想："可以这么说吧。我上大学时，曾和人合住过一套公寓，那里面老鼠极多，我就放了一只捕鼠夹。后来，我终于逮住了一只，可我无法把它从夹子上弄下来，因为我觉得这只死老鼠太可怜、太不幸了。因此，我就将死老鼠连同鼠

夹一起埋在后院里。后来，我决定情愿和老鼠住在一起，也不愿杀死它们了。”

“但你是吃荤的，”亨特指出，“因此，这同你亲手杀生实际上是一回事。”

德克森有点生气地说：“你的论点漏了重要的一点，那就是对生命最基本的尊重。我们同动物有相同之处。你没看见吗？”

“克兰有一个理论，你会觉得很有意思，”亨特还是坚持自己的观点。“他认为无论是真正的还是假设的生物血缘关系都同你对生命的尊重毫不相干。实际上，你不愿杀生，只是因为动物不愿死。它会喊叫、挣扎，一副可怜相，它苦苦哀求你不要杀它。还有，恰恰是你的心灵，而不是你的躯体，才听见动物的哀求。”

她看着他，满腹疑窦。

亨特把钱放在餐桌上，把椅子往后一推，说道：“你跟我来。”

半小时以后，德克森和亨特一起来到克兰家的大门前，大门自动开了，他们走到前门，门上没有锁，手一碰就开了。

德克森跟着亨特来到设在地下室的实验室里，里面有几十个柜子，他打开其中的一个，从里面取出一只银制的大甲虫，上面有彩色的指示灯，平滑的表面上还有若干机械的凸起物。他抓起甲虫，让她看它的底部有三个橡皮轮子，还刻着这样几个字：三号兽。

他把甲虫放在瓷砖的地板上，打开它下腹部的一个开关。伴随着轻微的哼哼声，甲虫开始在地板上爬来爬去，像是在寻找着什么。它停了一会儿，接着又爬向靠近墙脚处的电源。它在电源插座前停下来，从它金属的身体里伸出一双爪子，探进插座。顿时，甲虫身上的一些指示灯开始发出绿光，紧接着发出

一阵如猫一样的叫声。

德克森饶有兴趣地打量着这个精巧的玩艺儿："一个机械动物，很灵巧，这里面有什么奥妙？"

亨特从旁边的一条长凳上拿起一把锤子，递给她："我要你杀了它。"

"你在说些什么呀？"德克森略有警觉，"我为什么要杀……砸碎那个……那个机器？"她后退一步，不愿接过锤子。

"这只不过是一次实验，"亨特说，"几年前我按照克兰的命令也这样试过一次，发现很有收获。"

"你有什么收获？"

"我发现了生死的意义。"

德克森的目光中充满疑惑。

"这个'野兽'没有伤害你的防卫系统，"他向她保证说。"只要你在追它时，别碰倒东西就行。"说着，他把锤子递给她。

她试着跨前一步，接过锤子，斜眼瞧着这个奇怪的机器，此刻它正在电源插座前，发出低低的叫声。她走向前去，弯腰举起锤子，突然，锤子在半空中停住了："但……它在吃东西。"她转脸对亨特说。

他哈哈大笑起来。她生气了，双手高举锤子，重重地砸了下去。

然而，随着一声令人害怕的尖叫，甲虫从插座那里抽身向后退去。锤子重重地砸在地板上，打在甲虫爬过的那块瓷砖上，瓷砖被砸得坑坑洼洼。

德克森抬眼看到亨特在放声大笑。甲虫向后退了两米，停住了，睁眼瞅着她。不，她认为它并没有在瞅她。德克森胸中生起一股无名火，她紧握锤子，小心翼翼地走向甲虫。甲虫又后退几步，身上的两盏红色指示灯以接近人脑 α 波的频率闪着一强一

弱的光。德克森抡起手中的锤子，使劲地砸了下去，没打中……

十分钟后，她回到亨特面前，累得满脸通红，气喘吁吁。在追赶甲虫时，她的身体撞了好几处，碰得很疼；当甲虫躲在一条长凳下时，她俯身去砸，不料头碰在凳子上，疼得很。“这就像逮一只大老鼠！它身上讨厌的电池什么时候才能用完？”

亨特抬手看了看表：“要是你追得它来回跑，我想电池最多还能用半个小时。”他指着躲在另一条长凳下的甲虫，此刻它又找到了另一个插座，正在那里摆弄。“不过，还有一个更简单的办法。”

“我想试试。”

“你把锤子放下，把甲虫拿起来。”

“就……拿起来？”

“对。它只从自己的观点出发来判断危险，它只认铁的锤头。它的程序不对不持武器的原生体进行防备。”

她放下锤子，慢慢靠近甲虫，它没有反应。叫声也停止了，指示灯闪着微弱的光。德克森伸出手来，小心翼翼摸着了甲虫，它的身体在微微震动。她用双手战战兢兢地抓起甲虫。这时，它身上的指示灯变成了绿色，通过它微微发热的金属外壳，她能够觉出里面发动机的声音。

“现在我该怎样对付这个愚蠢的家伙？”她气呼呼地问。

“很简单，你将它背朝下放在长凳上。这样的话，它就一筹莫展了。你爱怎么砸就怎么砸。”

“它要是不像人，我就敢砸，”她口中喃喃自语，照亨特的建议办了，非得让事情有个结果。

德克森把甲虫翻过来，放在凳子上，这时它身上的指示灯又发出了红光。肚子上的轮子空转了一会儿就停住了。德克森操起锤子，迅速举过头顶，瞄准陷入困境的机器一锤砸下去，

打得它失去了平衡，一只轮子飞向空中。被砸坏的轮子发出一阵金属摩擦的声响，甲虫的整个身体发狂似地转动着。紧接着，从它的腹下又传来一阵声响，甲虫终于停止了转动，身上的指示灯闪着悲伤的光。

德克森紧抿双唇，高举铁锤，准备作最后一次重击。就在她下手的一瞬间，甲虫忽然发出一声像婴儿啼哭般的呜咽。她双手一松，铁锤落地，身体不由自主地后退几步。甲虫身下流出鲜血般的润滑剂，甲虫的四周被染红了。她睁眼盯着亨特，不禁毛骨悚然："这是……这是……"

"这不过是架机器罢了，"亨特说，现在表情忽然严肃起来了，"瞧瞧这些，都一样，这是它进化的先驱。"他指了指周围各种各样的机器，它们正默默无语，虎视眈眈地看着他们俩。"甲虫跟它们不一样，它能察觉自己的厄运，并呼叫求援。"

"去把甲虫身上的马达关了，"她断然说道。

亨特走过去，想把那个小小的马达关掉。"恐怕你已将它砸烂了。"他从地上拣起锤子，"再给它致命一锤，怎么样？"

德克森倒退几步，使劲摇头。"你难道不能修复……"话音未落，就传来一声金属撞击的巨响。她又吓得退缩几步，转过脸去，哭叫声终于停止了，他们默默地离开了地下室。

D · R · 霍夫施塔特

反　　思

贾森·亨特曾说过这样一句话："我们很难始终知道谁或是什么会有感情。"这正是这段选文的关键所在。李·德克森刚

开始时认定自我繁殖能力是生物的根本特征，亨特便很快向她指出，无生命的机器也能自动汇编程序。还有，像微生物、病毒这样的东西体内也有自我复制的指令。它们有灵魂吗？很难说！

然后，德克森认为感情是区别动物和机器的分水岭。为了将这个问题弄个水落石出，作者绞尽脑汁，详尽讨论了感情器官，试图使你相信存在着机械的和金属的感情，这看来肯定有语词矛盾。这也许是下意识要求的表现。他在文章中用了像“铝制的甲虫”、“轻微的哼哼声”、“害怕的尖叫”、“瞅着她”、“陷入困境的机器”、“发狂似地转动着”、“闪着悲伤的光”，诸如此类的话。所有这些描写已经让人觉得头晕目眩了，而最后一个形容更生动得无以复加，作者在描写甲虫被砸碎时，他写道，“甲虫身下流出鲜血般的润滑剂，甲虫的四周被染红了，”从它身上(还是从他身上?)传来“婴儿啼哭般的呜咽”。

作者在这里所作的描写很富挑衅性，读者不免要上当受骗。你会觉得自己被愚弄了，但是，愤慨之余本能的同情心却油然而生。有的人连放水冲走一只蚂蚁都下不了手。有的人却每天用活金鱼喂宠爱的猫。这里的区别又何在呢？究竟哪些东西是神圣的，而哪些东西又是无足轻重的呢？

我们中间很少有人是吃素食的，甚至从来没有想过这辈子是不是从吃荤转为吃素。这是不是因为我们觉得杀牛、杀猪等就无所谓？不是，我们都不愿意在吃排骨时联想到盘中装的会是一堆动物的尸骨。我们的语言形成了一些隐含的用法，并逐渐构成了一套规矩，以维持双重的标准。食肉的本质实际上跟性交、排便的本质是一回事，但我们在谈到这些事情时都比较含蓄，用了许多委婉的说法和隐喻，比方说，像“炸牛排”、“作爱”、“去洗澡间”种种说法。我们也感觉到屠宰场里在残杀生

灵，但我们在大嚼烩炙时都不愿联想到此情此景。

机器也一样。名叫象棋挑战者7号的机器人会下一手好棋，它在“考虑”下一步棋时，身上的红灯会轻松地闪亮；还有你小时候对它视如珍宝的精巧的玩具小熊。消灭其中的哪一个更容易下手？到底是什么在揪着你的心？这些机器使人联想起小巧玲珑、天真无邪、易受伤害的形象。

人类的情感很丰富，但同时又对灵魂的属性横加挑剔。纳粹分子是怎样为自己可以毫不顾忌地杀戮犹太人开脱的？为什么美国人在越南战争中能够肆无忌惮地“蹂躏黄种人”呢？似乎某一种感情——在这里是爱国热情——就像一个阀门，它能够控制其他一些感情，而这些感情会使我们以为我们的牺牲品实际上就是我们自己。

我们在不同程度上都是万物有灵论者。我们有些人认为车也有“人格”，还有一些人认为打字机或是玩具“是有生命的”，是有“灵魂”的。有些纸张我们很难付之一炬，因为上面写着我们的名字，一旦付之一炬，我们的一部分就随之化为灰烬了。十分明显，我们赋予这些杂物的“灵魂”实际上是我们自己心灵中的形象。但是，如果真是这样的话，我们赋予亲朋好友的灵魂不也是我们心灵中的形象吗？

我们都有移情活动，有时捉摸不透，而有时却一目了然，这取决于我们当时的心境，也取决于产生移情作用的不同的刺激物。有时候，随便的几句话就能打中要害，使我们的心肠顿时变软。而有时候我们却无动于衷，麻木不仁、冷若冰霜。

在这篇文章中，甲虫对死亡的抗拒深深地打动了德克森的心，也打动了我们的心。我们看到小小的甲虫在为自己的生命而搏斗，若用英国诗人迪伦·托马斯的诗句来形容，就是“因光明的消逝”而怒，拒绝“老老实实地走进那个黑夜而歇息”。作者

假设甲虫已经预感到了自己的厄运，这一笔最扣人心弦。它使我们联想起实验室中那些不幸的动物，它们被任意挑选，任意宰割。它们在杀身之祸临头之时，浑身颤抖。

一个身体到什么时候才有灵魂?在这篇充满激情的文章里，我们看到“灵魂”的出现并不是定义明确的内部状态的作用，而只是我们心理投射能力的作用。令人奇怪的是，这正是彻头彻尾的行为主义观点！我们对内部机制从不过问，而是将它归因于外部行为。图林曾对“心灵探寻”作过实验研究，即图林实验，而上述观点为证明图林实验提供了一种奇怪的辩护。



精神

艾伦·惠利斯

我们的产生如同一条长绳上微微发粗的顶端。细胞繁殖，变成赘生物，接着呈现出人形。长绳的顶端而今包藏在身内，受着保护，免遭破坏。我们的使命就是将它传宗接代。我们不过昙花一现，歌舞一时，留下一些刻在石碑上的记忆，然后，生命凋谢，我们消逝了。长绳的顶端如今深藏在我们子孙的体内，与我们紧密相联，然后通向远不可测的盘古。在这条长绳上曾出现过无数次增厚加粗，也有兴衰，同我们现在一样。一切都消逝，唯有细胞之绳长存。在生命长河中，导致创造新结构的东西，并不是昙花一现的赘生物，而是深埋于这条长绳之内的遗传机制。

我们是精神的载体。我们既不知道这是怎么回事，也不知道为何，也不知道往何处去。在我们的肩头，在我们的眼中，在痛苦的手中，通过不明的王国，走向未知的与不可知的未来，繁殖不断，我们承受了精神的重压。虽然精神以我们为存在之

依托，但我们却对它一无所知。我们以每一次心跳将它一寸寸地引向前方，把双手和头脑的劳作奉献给它。我们倒下了，将它传给我们的后代，我们的白骨被埋葬了，被丢失了，被遗忘了；而精神却代代相传，不断扩充，不断丰富，更加奇特，更加复杂。

我们被利用了。我们难道不该知道在为谁服务吗？我们将愚忠献给了谁？提出这样的问题意味着什么？除了我们所拥有的东西之外我们还能希望什么？什么是精神？

雅克·莫诺这样写道，一条河或是一块岩石，“我们知道或者相信是由物质力量的自由作用形成的，对这种自由作用我们无法提出任何设计、‘计划’或意图。也就是说，如果我们接受科学方法的根本前提，即，自然是客观的，而不是想象的，那是无法办到的。”

这个基本前提具有强大的魅力。因为，我们还记得这样的情形：就在几代人以前，那个时候，对立之物似乎一目了然，岩石想要崩裂，河流想要歌唱或咆哮；任性的精神在宇宙中遨游，用奇思异想支配着大自然。人类在理解和控制上的知识增加，是因为采用了这样一个观点，即自然之物和事件是没有目的或意图的。岩石并不想要干任何事，火山并不追求某个目标，河流并不希望寻找大海，而风也不曾追寻过自己的归宿。

但还有另外一种观点。原始的泛灵论并不是取代科学客观性的唯一选择。这种客观性在我们习惯思维的有限时空内也许是有效的，但在一个更大的时空内也许就不真实了。光以直线运动，不受周围物质影响，这种看法在日常生活中是可靠的，但在描述遥远的星系时就不那么有效了。同样，大自然是“外在的”，是没有目的的，这种观点在有限的人生中是正确的，但在永恒的旷野上也许会使我们误入歧途。

精神轻升，物质浊沉。精神如一团火，迈着舞蹈家般轻盈的步子来到世上。从虚无中，它创造了像上帝那样的东西，它即是上帝。精神诞生于某个起点，而这个起点又会是更早的起点的终点。如果我们追忆往事，就会发现自己站在原始的雾霭中，精神只是不安定的原子，一种颤动的东西，这东西不愿在沉寂和寒冷中缄默。

物质希望宇宙均匀扩散，静止而完美。精神希望有世界、天堂和地狱，旋转和冲突，有一个炽热的太阳来驱散黑暗，显露善恶；精神希望有思维、记忆、欲望，希望建造一条形式的长梯，不断变得复杂，不断包罗万象，通往不断上升、不断变幻的天堂，而一旦到达天堂，天堂又立刻化为通往更遥远的天堂的道路；最后……但是没有最后，因为精神永远向上，从不停息，尽管中途徘徊、盘旋、骤降，但它总是永远向上，毫不怜悯地用低级形式来创造高级形式，继续走向更深刻的内心世界，走向更大的自由。

粒子活了。精神跳离了物质，而物质永远要拖住精神，拉下它，让它静止下来。微生物在温暖的海洋里蠕动。这些一时带着追求精神的小生物变得更加复杂。它们聚在一起，互相触摸，精神开始创造爱。它们触碰着，有东西从它们身边掠过。它们死亡、死亡、死亡，不断地死亡。谁了解那些隐没在历史长河中的小生物呢？谁又了解在古老的海中跳动的银鱼？谁能听见从未听见过的拍岸的浪涛？谁又看见旷野上的野兔，成群的野鼠？它们死亡、死亡、死亡，但它们互相触摸，有东西从它们身边掠过。精神腾跃，又去创造新的躯体，从不间断，创造更加复杂的容器，把精神载向未来，将更丰富的精神留给后代。

病毒变成细菌，又变成水藻，再变成蕨草。精神的力量使岩石崩裂，使洋松猛长。变形虫伸开柔软的触角，不断地摸索着世界，想更加了解世界，将世界装入心中；它不断长大，更加好奇，精神更加丰富。变形虫变成柔鱼，又变成鱼；扭动变成游动，又变成爬行；鱼变成蜓蚰，又变成蜥蜴；爬行变成步行，变成奔跑，又变成飞行。生物互相接触，精神在其中跃出。生物向性变成嗅觉，变成迷恋，变成欲望，又变成情爱。从蜥蜴到狐狸，到猴，到人，一句话，我们聚在一起，触摸、死亡，为精神服务着，却不知道已将精神代代相传。精神的双翼愈加丰满，精神的飞跃愈加巨大。我们爱恋遥远的情人，怀念早已入土的亲友。

* * *

"人是精神的载体，"埃里克·赫勒这样写道，"……精神这个旅行者，穿过人的大陆，招呼人类灵魂随它而去，走向纯精神的目的地。"

若俯地观察，我们发现精神之路弯弯曲曲，恰似菲菲黑森林中蜗牛闪亮的爬迹，倚天鸟瞰，我们看见弯弯曲曲的支流汇入滚滚向前的大江大河。人类已爬上一座高山，在山上可以回首远眺。千百年来眼前的景象一直十分清晰，再过几千年，这景象尽管会有些模糊不清，但仍然依稀可辨。在我们最后一次旅行的变化无常的路口，伸展着一条通向远方的笔直的金光大道。人类无法使其产生和终结，人类永远行走在这条路上，披荆斩棘，逢山开路，遇水架桥。我们走过的道路是属于谁的？它不属我们人类。因为它上面有人类的第一个脚印；它也不属于生命，因为在生命尚未出现以前，这条道路便已存在。

精神是位旅行家，现正穿越人类的王国。我们没有创造精神，也不拥有精神，也无法定义精神，我们只是精神的载体。我们把尚在襁褓中的精神背在肩上，背着它渡过我们的生命旅程，将它代代相传，在这同时，它或是变得丰富了，或是变得萎缩了。精神是旅行家，人是旅行工具。

精神创造，精神毁灭。没有毁灭的创造是不可能的；没有创造的毁灭靠过去的创造维持生命，退化为物质，趋向静止。精神所创造的东西大于它所毁灭的东西(但并不是每时每地都如此，因此那些支流和小径也会逆转。有时物质追求死寂的愿望也会得到满足)，但不管怎么说，创造的优势保证了整个进程的稳固性。

从原始的物质雾霭到复杂的星系和精确的太阳系，从熔岩到充满空气、土地和水的地球，从沉重不堪到生命的展翅欲飞，从感觉到知觉，从记忆到意识——人类现在手持一面明镜，精神在其中看清了自己。江河中浪头回扑，旋涡四起，江河飞泻、消失、重现，又滚滚向前，前进的主流是形式生长、意识增强，物质变为心灵又变为意识。人和自然的和谐寄寓于生命的旅程沿古老的行程走向愈来愈大的自由和愈来愈强的意识。

D·R·霍夫施塔特

反　　思

精神病学家艾伦·惠利斯用诗一般的语言向我们描述了一种怪异的、使人迷惑的观点，即现代科学使人类找到了自己在物质世界里的定点。许多科学家，且不包括人文学家，发现这

个观点难以接受，他们试图寻找某种也许是捉摸不定的精神本质，来作为区别生命、尤其是人和其他非生物的标准。生命如何从原子中产生？

惠利斯的“精神”并非这样一种本质。他以此来描绘似乎有目的的进化过程，好像在这背后有一股力量在指引它前进。如果真有什么目的，那么它就是理查德·道金斯在下一篇文章中明确阐述的那种东西：稳定复制者生存。在序言里道金斯直言不讳：“我们是生存机器——是盲目地编入了保护被称为基因的分子程序的自控机。这个真理至今仍使我吃惊不已。尽管我早已知道，可我从来没有完全适应过。我有一个希望，也许能使别人也感到吃惊。”

10 自私的基因和自私的米姆

理查德·道金斯

自私的基因

起初事物都是简单的。连解释单个的宇宙如何发生这样的问题都很困难。我假设谁都同意，若要解释复杂的次序——生命，或一个创造生命的东西——如何突然产生，而且样样功能具备，那将还要困难。达尔文的自然选择的进化论是令人满意的，因为它表明了由简单向复杂进化的过程，表明了无序原子如何重新排列，形成更加复杂的序列，最后创造出了人的过程。达尔文提供了迄今为止唯一令人满意的解决有关我们存在问题的答案。我打算用一种比习惯方法更基本的方法来解释这个伟大的理论，而且以进化本身尚未开始的年代为起点。

达尔文的“适者生存”实质上只是“稳定者生存”这个更普遍规律的一个特例。宇宙间居住着稳定的事物。一件稳定的事物是一些原子的集合，它稳固、普遍，因此有一个名称。它可以

是一些原子的特殊集合，如马特峰①因为它在世界上存在了这么久，值得有一个名称。它也可以是一类实体，比方说雨点，它们由于以群体出现而获取一个集合名称，尽管一滴雨点的寿命很短。在我们周围的、我们觉得需要解释的事物，像岩石、星系、海浪等，多多少少都是原子的稳定模式。肥皂泡一般都是圆的，因为这是充满气体的薄膜的稳定外形。在宇宙飞船里，水也呈稳定的球形，而在地球上，由于地球引力，水的稳定表面是平坦的。盐晶体一般是立方体，因为这是氯和钠结合的稳定形式。在太阳里，最简单的原子即氢原子聚变为氦原子，因为在那里氦化形式更加稳定。在宇宙各处恒星内通过"大爆炸"形成了更复杂的原子，按照目前流行的理论，"大爆炸"便是宇宙的源泉。这就是我们这个世界的元素的来历。

有时候，原子在化学反应中连接在一起，形成了分子，分子多多少少是稳定的。如此形成的分子会很庞大。像宝石这样的大结晶体也能被认为是一个单分子，它不仅稳定，而且也很简单，因为它的内部原子结构是无限重复的。在现代生物内，还有其他高度复杂的大分子，其复杂程度表现在若干层次上。人的血液内的血红蛋白就是一个典型的蛋白分子。它由一系列较小的分子和氨基酸所构成，每一个都含有几十个排列整齐的原子。在血红蛋白分子内，有574个氨基酸分子。这些氨基酸分子排列成4条链，彼此缠绕交错，形成一个极其复杂的球状三维结构。血红蛋白分子模型看上去像一簇多刺高灌丛，但又不同于真的灌丛，它不是杂乱无章的，而是一个确定不变的结构，在一般人体内反复重复达亿万次，而不出半点差错。像血红蛋白这样的蛋白分子多刺高灌丛是稳定的，因为由同样氨基

① 阿尔卑斯山之一座高峰，在意大利与瑞士边境。——译者

酸序列构成的两条链，如两股溪流汇入同样的三维螺旋结构中，血红蛋白多刺高灌丛会以每秒亿万次的高频在人体内长成它们“喜爱”的形状，同时，有些血红蛋白也以同样的频率被消灭了。

血红蛋白是一个现代分子，用来说明这样一个原则，即原子倾向于组成稳定的结构。这里有必要指出，在地球上的生命出现以前，某些最基本的分子进化可以由普通的物理和化学变化所引起。没有必要去考虑设计或目的之类的东西。如果一组原子在能量的作用下形成了一个稳定的结构，该结构就会一直存在下去。自然选择的最初形式无非是选择稳定形式，摒弃非稳定形式。这一观点并不神秘。出现这一情况也是必然的。

当然，这并不是说，你就可以原封不动地依照这些原则来解释像人类这样的复杂机体的存在。你不可能拿若干个原子，将它们放在某个容器内，然后用外力摇晃，直到那些原子碰巧形成了所需的结构。接着，亚当就从瓶口内跳了出来！也许你能够这样用几十个原子来形成一个分子，可一个人是由亿亿万万个原子所构成的。你要是想造一个人，你就得用生化鸡尾酒混和器不停地搅动，搅动的时间之久使整个宇宙史在你看来也只不过是一眨眼的功夫，即使是这样漫长，你也不一定会如愿以偿。达尔文的进化论从总体上解决了这个问题。就在分子缓慢形成的故事中断的地方，达尔文的理论接下去了。

我所要作的对于生命起源的解释无疑是纯理论的。因为谁也没有看到过真相。对此有一些互相对立的理论，但它们都有一些相同的特征。我将要作的简化了的解释也许离真相不会太远。

我们不知道在生命出现以前，地球上有哪些化学原料是充分的。但像水、二氧化碳、甲烷和氨这样的物质是有的，这些

简单的化合物至少在太阳系的其他星球上也存在。化学家们一直在试验模拟早期地球的化学环境。他们将这些简单物质放入一个试瓶内，并提供如紫外线或电花之类的能源，这些都是原始闪电的人工模拟。几个星期以后，试瓶内会出现有趣的现象：试瓶内出现了稀薄的棕色汤，汤内溶有大量的分子。这些分子比原先放入的分子结构要复杂得多。特别值得一提的是发现了作为蛋白质构成的氨基酸，这是两大类生物分子中的一类。在这些实验以前，人们认为自然出现的氨基酸是生命出现的征兆。如果在哪儿发现了氨基酸，比方说在火星上，那么，火星上出现生命就为时不远了。而现在，氨基酸的出现只说明了在大气中存在若干简单气体以及火山、光、雷雨气候。最近，在对生命出现以前的地球上的化学环境的人工模拟中，又发现了被称为嘌呤和嘧啶的有机物。这些物质是遗传分子 DNA 的构成。

与这些实验相类似的过程一定产生了“原始液”，生物学家和化学家相信这种“原始液”在三、四十亿年以前构成了地球上的海洋。也许在岸边逐渐干涸的浮垢内，或悬垂的小水滴内，这些有机物质就近聚集起来了。再加上如来自太阳的紫外线的能源的影响，它们聚成了更大的分子。现在，较大的有机分子生存时间较短，而不被我们注意，因为它们很快被细菌或是其他的生物吸收和破坏了。但是细菌和别的生物在当时还没有出现，这样，这些大有机分子能够平平安安地飘浮在愈来愈浓厚的“原始液”内。

在某一时刻，由于偶然的因素，一个特别重要的分子形成了。我们称之为复制者。它并不一定就是周围分子中最大或最复杂的一个，但它具有能复制自己的特殊本能。看起来这类偶然事故很罕见。的确是这样。这样的事极难发生。在人的一生

中，不易发生的事实际上被认为是不可能发生的事。这就是为什么你永远不会在赌场上获大奖的原因。但是，用我们人类对于什么可能发生以及什么不可能发生的估计来考察几亿、几十亿年以前的事，我们就感到无所适从了。如果你每星期都往赌盘上掷几个筹子，这样做上一亿年，你保险中好几回头奖。

实际上，复制自身的分子并不像最初那样难以想象，它曾是偶然的产物。将这个复制者当成模盘或模板，将这个复制者看作一个大分子，这个分子具有一条由各种基本构成分子构成的复合链。这些基本构成分子在复制者周围的液体里比比皆是。现在假定每一个基本构成分子对同类都有亲和性。这样，每当一个外来构成分子靠近复制者的那部分与其有亲和性的分子时，就会粘上去。以此种方式构成的基本构成分子会自动形成模仿复制者结构的序列。很容易设想，它们聚集起来，形成一个稳定的链，与原来那个复制者的内部序列一模一样。这种过程能够一层一层地反复进行。晶体就是这样形成的。另一方面，这两条链也有可能分裂，如果这样的话，就会出现两个复制者，它们各自又都能继续复制自身。

还有另外一种更复杂的可能性，每一个基本构成分子与自己的同类并无亲和性，而与另外某一个种类却有互相亲和性。这样，复制者的作用不是直接铸造一个一模一样的复制品，而是充当一张“负片”，由此“负片”再重制一张原件的“正片”。对于我们来说，原件的复制过程是正片负片，还是正片正片，是无关大局的，尽管有一点值得重提，即复制者的现代形式DNA分子使用了正片负片的复制过程。对我们来说，重要的是，一种新的“稳定性”突然降临于世。起先，某种复杂分子在原始液内不大可能大量产生，这是由于每一个分子都依赖于碰巧形成某一种稳定结构的基本构成分子。复制者一出现，它就迅速在

海洋中四处传播它的复制品，以至较小的基本构成分子越来越少，其他较大分子的形式越来越罕见。

这样，看上去我们有了相当数量的复制品。但现在我们必须指出任何复制过程中都具有的一个重要特点：它不是完美无缺的，而是会发生错误的。我希望此书不会有印刷错误，但你只要仔细检查，你就会发现有一、两处错误。这些错误不会严重歪曲句子的意思，因为它们是“第一代”错误。但是，想象一下，印刷术还没发明的年代，像《福音书》这样的书是手抄的。无论你多么小心，错误是不可避免的，有时会错得面目全非。如果所有的手抄本都来自于同一个权威原本，那么，意思不会被歪曲得太多。但是，如果一个抄本抄自另一个抄本，而这另一个抄本又抄自再一个抄本的话，那出现的错误将会越来越多，并越来越严重。我们一般都认为错误的抄本是劣制品，譬如说文件，很难设想会有将错误视作改进的例子。我假定《旧约全书》的希腊本译者由于翻译的错误而惹出了一个大麻烦。他们将希伯来语的“年轻妇女”错译成希腊语“处女”，这样，《圣经》中就出现这样一句预言：“看啊，处女将怀孕并生一个儿子……”但是，我们会发现，生物学中的复制错误却真的能导致改进，引起进化。对于不断进化的生命来说，出现一些错误是不可缺少的。最早的复制分子如何确切地复制的情况，我们尚不清楚。原始分子的现代传人——DNA分子，同大多数高度诚实的人类复制过程相比较，要坚贞不渝得多，但即使是它，也不时要犯错误，正是由于这些错误才使进化有可能发生。也许最初的复制者更是变化无常，但不管怎样，我们可以肯定出现了错误，这些错误越积越多。

随着复制错误的出现和扩展，“原始液”里就不光有相同的复制品，而且有若干不同的复制分子，它们都来自同一个祖先。

是不是有些品种的队伍会更庞大？几乎肯定会这样。有些品种天生更稳定。有些分子一俟形成就不易再分裂。这类分子的数目就会相对庞大，这不仅是由于“长寿”的直接逻辑结果，也是由于它有充足的时间复制自己，长寿的复制者因此会更多，要是别的条件都一样，分子的“进化倾向”就会向长寿发展。

但也许别的条件并不一样，复制者品种在传播种系中还有一个更重要的特征，这就是复制的速度，或者说是“生育力”。假定复制分子A型平均每星期复制一次，而B型则平均每小时复制一次，这样，不难看出，B型分子数量就会很快超过A型，尽管如果A型的寿命要比B型长得多。这样，也许就会出现朝分子的“高产”发展的“进化倾向”。本来必定会形成的复制分子的第三个特征是复制的精确性。假设X型分子和Y型分子有同样的寿命以及同样的“生育力”，但X型每复制10个就出现1个错误，而Y型每复制100个才出现1个，很明显，Y型分子的数量就会多些。而X型分子不仅失去了因复制错误而诞生的“孩子”，也失去了他们所有现存或将来的后代。

如果你对进化论略知一二，会发现这最后一种说法有一点小矛盾。认为复制错误是进化发生的重要先决条件的观点，是否与认为自然选择喜欢复制的高度忠实的说法相违背？答案是：尽管模糊地说，进化看上去是件“好事情”，特别是人类是进化的产品，但实际没有什么东西“要求”进化。不管愿意不愿意，进化发生了，尽管复制者（现在的基因）千方百计阻止过进化的发生。雅克·莫诺在他的赫伯特·斯宾塞讲座中，对这一点作了很好的说明，在这样说明之前，他不无讽刺地评论道：“关于进化论的另一个奇怪的问题是，每个人都自以为懂得进化论！”

现在让我们再来看“原始液”，这里肯定已经生存着各种稳定的分子。这里说的“稳定”有三种情况，或者是个体分子生存

时间较长，或者是复制速度较快，或者是复制的精确度较高。而朝这三种稳定形式发展的进化倾向指的是以下意思：如果你分两次采样“原始液”，后采的样会含有更多具有较长寿命、较高生育力和较忠实复制品的分子品种。在生物学家讨论生物时，他所说的进化正是指以上这种情况，机制都一样——自然选择。

我们是否应该把最初的复制分子说成是“活的”？谁也不会对这感兴趣。我会对你说：“达尔文是人类历史上最伟大的人。”而你会说：“不对，应该是牛顿。”但是，我希望我们不要在这上面多费口舌。无论我们采用什么方法解决争端，结果不会给实质问题带来什么影响。不管我们怎样评价牛顿和达尔文，说他们伟大也好不伟大也好，他们的生活和成就并不会因此而有丝毫的改变。同样的道理，不论我们说复制分子是否是“活的”，它们的进化历史也许正如所表达的那样，并不因此受到丝毫的影响。语词不过是我们生活的工具，许多人由于不明白这个道理，因此造成生活中的许多磨难。我们的词典里出现了“活的”这样一个词，这并不意味着现实世界里就非得有十分确切的指代物。无论我们说早期的复制者还活着也好，已经死了也好，它们还是生命的源泉，是我们的祖先。

争论的下一个重要问题是关于竞争的，达尔文自己对此十分看重（但他当时讨论的是动植物，而不是分子）。原始液并没有能力维持在不断递增的无数复制分子的需要。因为地球的体积是有限的，此外，还有其他一些重要的限制因素。我们在前面已经说过，复制者作为模盘和模板，生活在充满基本构成分子的原始液之中，这样它就必定能够制造拷贝。但是，随着复制者的增多，基本构成分子被消耗得太快，变得越来越少，并且越来越珍贵。各类的复制者必定为此而竞争。我们已经考虑

过使有利的复制者数量增长的因素。而有些不利的品种则由于竞争而数量日益减少，最后灭绝。各类复制者为了生存进行斗争。当然，它们并不知道自己在进行残酷的竞争，或在为竞争而担忧，斗争毫无残酷感觉，其实没有任何感觉。但它们的确在斗争，任何导致新的更高稳定层次的复制错误，或者能减低对方稳定状态的新方法，都被自动保留下来了，并通过繁殖扩散开来。进步的过程是渐进的。增强自身稳定和削弱对方稳定的方法变得更复杂，更有效。有的复制者甚至"发现"了怎样通过化学途径去破坏对方的分子，怎样在对方遗弃的基本构成分子中复制自己。这些可以被称为最早的食肉动物，它们在获取食物的同时也消灭了竞争的对手。有的复制者也许发现了如何保护自己的方法，它们或者用化学手段，或者在自身周围筑起一道蛋白质的围墙。这也许就是第一个活细胞的来源。从这以后，复制者不仅开始存在，而且也开始为自己制造容器和器具。那些为自己制造了作为栖身之所的生存机器的复制者就生存下来了。最早的生存机器大概只是一层保护膜。但是，随着进化的发展，各类的复制者都制造出了更好、更有效的生存机器，这样，谋生越来越不易了。生存机器变得越来越大，越来越复杂，这个过程是渐次累进的。

复制者为了保证自己生命的延续，在不断地设法改进技术和手段。这个持续的进步过程是不是有个终结呢？改进的时间很充裕。在一千年的时间里会产生多么离奇的自我保存机器？四十亿年过去了，古老的复制者今在何方？它们没有消亡，因为它们是生存艺术的前辈。但是，你别想在汪洋大海中寻找它们的身影，它们早已结束了这种流浪的生活。而今，它们聚集在广阔的殖民地里，远远躲进外形笨重、庞大的机器人中，同外部世界已不相往来，只能通过转弯抹角的间接渠道与世界交

流、并通过遥控来操纵这个世界。它们就在你和我的体内，它们创造了我们，我们的肉体和灵魂，它们的保存是我们的存在的最终依据。这些复制者经过长途跋涉才走到这里。现在它们改名换姓，变成了基因，我们只不过是它们的生存机器而已。

*　　*　　*

从前，自然选择包括飘浮在原始液中的复制者的不同存在方式。现在，自然选择有利于那些善于制造生存机器的复制者，也就是有利于那些精于控制胚胎发展的基因。在这里，复制者同以前一样，仍旧无意识、无目的。互相竞争的分子之间自动进行选择，主要依据三个因素，它们是寿命、多产和复制精确性，这个过程依然像远古时代那样继续没有目标地进行下去，永不停息。基因不能看到将来。它们不会事先订计划。基因只是存在着，有的存在久些，仅此而已。但是，决定基因寿命和多产的性质就不那么简单了。它们并没有那么漫长的历史。

近年来——大约六亿年的样子吧——复制者在创造生存机器技术方面取得了十分突出的成就，它们已经能够造出肌肉、心脏和眼睛了（眼睛独自进化了好几次）。在这之前，复制者彻底改变了自己生活方式中的一些基本的特征，如果我们想继续进行论证，就必须先来谈谈这些基本特征。

现代复制者的第一个特征是它们的高度群集性。一个生存机器里面装载着几千几万个基因，而不是一个基因。基因高度密集，互相合作，不分你我，才构成了身体。一个特定的基因对身体的不同部位会产生不同的影响。身体的某一特定部位受着许多基因的影响，而其中每一个基因的影响都依赖于与其他基因的相互作用。有些基因在一群基因的活动中起着领袖的作用。例如，建筑蓝图的每一页都代表了建筑物的不同部位，每

一页蓝图只有同其他许多页相互参照才能有意义。

你也许会这样想，既然基因互相依赖，紧密相连，那么，为什么我们还要用“基因”这个词呢？为什么不用一个像“基因复合体”这样的集合名词呢？回答是，的确，这个想法在许多方面都不错。但是，如果从另外一个角度看问题，那么，我们把基因复合体分成分立的复制者或基因，也的确是有道理的。因为这里牵涉到性的问题。通过性繁殖，基因被混合、被搅乱了。这就是说，个人的身体只是装载寿命短暂的基因混和体的暂时容器。作为任何一个个体的基因混合体的寿命会很短暂，但是，基因本身能够活得很长。它们的道路永远一代一代地纵横交错。一个基因是一个独立的单位，它能够在许多代延续的个体身体中生存下来。

*　　*　　*

自然选择从最广义上说指的是实体的差别生存。有的实体活下来了，而有的却死了，但是，为了使这种有选择的死亡对世界产生影响，还必须满足另外一个条件。每一个实体是以众多的复制形式存在的。这中间至少有一部份实体具有生存的潜在能力，它们以复制的形式能活上许多年。小的基因单位具有这些属性；而个体、群体，乃至物种则没有。格雷戈尔·孟德尔曾证明在实际做法中我们能够将遗传单位当作不可分割的独立的粒子，这在当时是一个很重要的发现。现在我们知道这有点太简单了。即使是一个顺反子也偶尔能被分割。还有，同一染色体内的一对基因也不是完全独立的。我从古到今说了那么多，意在说明基因是一个单位，并在很大程度上接近不可分粒子状这个现象的概念。一个基因不是不可分的，但很少被分割。在任何个体的躯体内，它不是肯定存在，就是肯定不存在。一个基

因从祖辈传到儿孙辈，完好无损，中间经历了许多代，并没有同其他基因相混合。如果基因一直不停地互相交杂，那么，我们所理解的自然选择就不可能发生了。恰好达尔文在世的时候，这种说法就已经被证实了，达尔文为此而很担心，因为当时人们都认为遗传是一个交杂过程。孟德尔的发现当时已经发表了，他的理论本来能够解救达尔文，可后者对此却闻所未闻。当时孟德尔的理论鲜为人知，只有当孟德尔和达尔文都去世以后，人们才重新发现这个理论。孟德尔也许没有意识到自己理论的重大意义，要不然，他也许就写信告诉达尔文了。

基因粒子状的另一个特点是不会衰老。一个基因无论已活了一百万年还是只活了一百年，都不会死亡。它一代又一代地由一个身体传给另一个身体，用自己的方式和根据自己的目的操纵着身体，不知要在跨越过多少个身体之后，才开始衰老、消亡。

基因是不朽的，确切地说它们几乎是不朽的遗传实体。我们，世上个别的生存机器，可望再多活几十年。而基因的寿命却不是十年的问题，而是几千年，几百万年。

* * *

生存机器在开始时是基因的贮存器，只提供一堵墙以免遭竞争对手的化学性袭击和偶有发生的分子攻击。在早期，它们以"食"原始液中的有机分子为生，这些分子在当时唾手可得。然而需要有许多世纪的太阳能作用才慢慢地产生了这些有机分子，于是它们很快被消耗尽了，这种舒适生活也随之结束了。生存机器的一个重要分支，现在叫植物，开始直接利用阳光从简单的分子自制更加复杂的分子，以快得多的速度重新进行原始液的合成过程，另一个分支，现在叫动物，"发现"该如何利

用植物的化学成果，它们吃植物，或吃其他动物。生存机器这两个重要的分支不断改进着技艺，以提高生存的效率，因此，新的生活方式不断出现。这两个分支又发生了新的分化，新分化出来的亚分支在生存技艺方面各有特点：它们分别生活在海洋里、陆地上、天空中、地下、树上或是在其他生物的体内。由于这种再分化，种类繁多的动植物产生了，这才有了今天的丰富多采的世界。

不论动物还是植物都进化成了多细胞的身体，基因分布在每一个细胞之中，身体是基因完整的拷贝。我们不知道这是在何时，为何故发生的，也不知道各自费了多少波折。有些人把身体比作细胞的殖民地。我认为身体应该是基因的殖民地，而细胞只是基因的化学工业的颇为方便的工作场所。

尽管身体只是基因的殖民地而已，但我们不能否认它们拥有自己的独特性。一个动物是一个协调的整体，是一个独立的单位。我从主观上觉得自已是一个独立的整体，而不是一块殖民地。这样想是合乎情理的。选择是欢迎那些与其他基因合作的基因的。在为了争夺稀少的资源的残酷竞争中，在捕食其他生存机器的无情战争中，在防止被他人吞食的垂死挣扎中，一个独立的身体的中枢协调性能越好，它的生存能力就越强。现在，基因之间的共同进化程度可以说是到了这样的地步，一个独立的生存机器其内部基因配合相当默契，我们已很难发现这竟是无数基因在其中密切合作的结果。实际上有许多生物学家没有发现个中奥妙，他们不会同意我的看法。

* * *

生存机器行为最突出的特征之一就是它的明显的目的性。我说这句话不仅是想指出生存机器似乎是有意要帮助基因生存

下去，虽然它是这么回事。我想指出的是动物的目的性行为与人类具有许多相似之处，每当动物"寻找"食物、配偶或丢失的幼兽时，我们会情不自禁地认为它们与我们人类具有某些同样的主观感觉。这些主观感觉包括对某个物体的"欲望"，对所需物体的"心理想象"、某种"目标"以及"期待的目的"。凭着内省的经验，我们每一个人都很清楚，至少在一个现代生存机器身上，这种目的性已经进化成了我们称之为"意识"的东西。我不是哲学家，不敢冒昧深涉其中的涵义，但值得庆幸的是这个问题与眼下的目的无关，因为很明显机器的行为似乎是有目的的，至于它们是否真有意识，仍可讨论。这些机器一般都非常简单，在工程科学中，无意识但有目的的行为的原则是较为普遍的。最典型的例子是瓦特蒸汽机上的调节器。

与此相关的最基本的原则叫负反馈，负反馈有几种不同的形式。泛泛而论的话，情况大致如下。所谓的"目的机器"也就是似乎有目的地表现其行为的机器和物体，它们具有某种测量器件，用以测量现在状态与"期望"状态之间的距离。这两种状态之间距离越大，机器所作的功也就越多。于是，机器就会自动地想要缩短两者之间的距离，这就是被称为负反馈的原因；一旦实现了"期望"状态，机器就会停止工作。瓦特调节器实际上是两个球，由蒸汽发动机带动旋转。两个球各自都系着一条铰链臂。球转得越快，将铰链臂推向水平线的离心力就越大，但这种倾向又受到引力的阻碍。铰链臂联着控制发动机的汽门阀，一旦铰链臂接近水平线，汽门阀就会自动关闭。因此，如果发动机工作得太快，其中一部分蒸汽就会被阻截，其工作速度就会减慢。如果速度太慢，汽门阀会自动打开供应蒸汽，发动机转速又会加快。由于时间差过大，这种目的机器经常左右摇摆，因此，安装减少摇摆的辅助装置也是很重要的一个环节。

瓦特调节器的“期待”状态是一种特殊的转速。很明显，调节器并不是有意想达到这种状态的。我们可以将机器的“目标”简单地定义为它倾向于返回的那种状态。现代目的机器则使用了基本原则的外延，比如说负反馈，以表现更加复杂的“有生命般的”行为。例如，导弹似乎能主动地追寻目标，在射程之内它能跟踪目标，并能排除目标有意躲闪的干扰，有时候甚至能“预料”或“预测”目标的行踪。至于导弹如何跟踪追击这里就没有必要深究了。有一点必须指出，各种各样负反馈在导弹中起着作用，此外还有“前馈”作用以及其他原则，这些原则工程师都了解，并且现已发现它们在生物体内也起着广泛的作用。与意识相去甚远的东西什么也不必考虑了，即便是一个门外汉，虽然观察到导弹的有意的、有目的的行为，也会觉得难以相信导弹不是直接由人来控制的。

很多人都有这样一个错误观点：认为因为像导弹这样的机器是由有意识的人设计并制造的，所以它们必须受有意识的人的直接控制。这种错误观点还有另外一种说法，即“计算机并不真正会下棋，它只是贯彻人的指令”。必须搞清楚为什么这种观点是错误的，这至关重要，因为这影响到我们对“基因能够控制行为”这句话的理解，计算机下棋这个例子很有说服力，因此，我想扼要地说几句。

计算机的棋艺即使不能与人类象棋大师相提并论，也达到了优秀业余棋手的水平。严格地说，是程序达到了优秀业余棋手的水平，因为对弈程序放在任何一台计算机中都能运用自如。那么，设计程序的人起着什么作用呢？首先，他并不每时每刻都在像玩木偶那样控制计算机，那样做只是在骗人。设计者编完程序后，将它输入机内，就让计算机独立运算了。除非对手输入自己的棋招，人不再干涉。设计者是不是有可能预测所有

的棋招，为计算机设计一连串妙棋，以应付对手的一着一势呢？这绝不可能。因为对弈时棋势变幻无穷，要想一应俱全除非到了世界末日。同样道理，要设计出使计算机的每一着棋都属于赢棋的步骤的程序是不可能的。对弈时棋势花样层出不穷，连星系中的原子数量都比不上。在计算机下棋这个问题里面奥妙无穷，我们暂且谈到这里。事实上，这个问题是非常难的，因此，设计得再好的对弈程序也只能达到业余棋手的水平，这并不奇怪。

其实，设计者的真正作用就好比一位慈父教幼子对弈一般。他告诉计算机最基本的招势，教它的不是一成不变的教条，而是一应俱全的规则。他不是一板一眼地说“象走对角线”，而是说这样一些经过数学概括的话（尽管还要更简洁些），比如“通过给象的旧x坐标和y坐标加上同样的常数，尽管并不一定是同样的符号，象的新坐标可以从其旧坐标中得出”。接着他会输入某些用同样数学或逻辑语言写成的“建议”，这些语言相当于人们所说的下述一类话：“别让你的王毫无防卫”或“用马同时踏两个子”。细说起来也很引人入胜，但那样一来便离题太远了。重要的就是：在真正对弈时，计算机是独立运算的，并不能从它的主人那里获得任何帮助。程序设计者所能做到的只是使计算机事先尽可能获得最佳方法，在专门的知识与关于策略和技巧的暗示之间保持恰如其分的平衡。

基因也同样控制着它们的生存机器的行为，当然不是像手拉控制木偶的线那样进行直接控制，而是通过间接的方式，就像程序设计者控制计算机那样。基因所能做到的只是事先造就生存机器，然后就由它自己去了，而基因只能被动地存在于生存机器之内。为什么基因如此被动呢？为什么它们不在里面采取主动，控制机器呢？原因是它们没有能力，因为存在着时间差所造成的阻碍。我们不妨打个比喻来解释这一点。弗雷德·

霍伊尔和约翰·埃利奥特曾合著过一本科学幻想小说《仙女座》，同所有优秀科幻小说一样，这本书包含了一些十分有趣的科学思想。奇怪的是，似乎书中没有明确提及这些重要的科学思想。它留给读者去想象。我想在这里展开这些思想，作者不会介意吧。

在200光年之遥的仙女座里，存在着一个文明社会。这个社会希望把他们的文化传播给遥远的星球。使用何种办法最快呢？直接旅行是不可能的。因为上有光速从理论上给旅行速度作的限制，下有制造宇宙飞船的实际技术性困难。还有，宇宙之内是否真有许多值得一去的星球，你知道往哪个方向飞吗？通过无线电与其他星球联络，这个想法更好一点。因为，如果你有能力，可以把信号发往四面八方，而不仅是发往一个方向，这样，你有可能联络上许多星球(星球数量的增加相当于信号发射距离的二次幂)。无线电波以光速运动，这也就是说，来自仙女座的无线电信号要等200光年以后才能到达地球。在此，距离过长是个棘手的问题，因为双方绝对没有可能进行对话。要是真的进行对话，那么，发自地球的每一个连续的信号将由12代人陆续完成。即使我们对这一点忽略不计，那么，通过如此遥远的距离进行对话也简直是荒谬绝伦。

这个问题很快就会成为一个十分紧迫的问题：无线电波从地球发送到火星需要4分钟。这样，毫无疑问，太空人的说话习惯必须改变，他不能用简短的对话形式进行通话，而必须用长段的独白，与其说像对话倒不如说像写长信。另外还有一个例子。罗杰·佩恩曾经指出，海洋声学有一些特点。座头鲸能发出音量宏大的“叫声”，从理论上讲，如果它的位置在海洋的某一特定深度，那么，其“叫声”就能传遍全球。现在还不知道座头鲸之间是否真的相距如此遥远进行交流，如果是这样的话，

它们将碰到和火星上太空人同样的麻烦。声波在水中传播速度很慢，在大西洋里走一个来回差不多需要两个小时。我想这可以解释下述事实，座头鲸在大声“喊叫”一阵以后，会沉默整整 8 分钟，然后重又叫喊起来，间隔一般都是8分钟。

仙女座星文明社会也像座头鲸一样。由于信号发出以后，没有必要空等，他们就将要说的话编成一个长长的电文，反复地播放这段电文，每播放一次中间停顿数个月。然而，他们的电文与座头鲸大不一样。在这个电文中，包括了一个编成电码的指令，指示怎样建造一个大型的计算机以及编制它的程序。当然，这些指令不是用人类的语言写成的，但是一个有经验的译码专家能够破译任何一个密码，尤其是编这些指令的目的就是准备让人看懂的。地球上的乔德雷尔·班克的射电望远镜收到了这个信号，最终破译了它，然后，按照指令建造了计算机并输入了指定的程序。对于人类来说，其结果几乎是灾难性的，因为仙女座那个文明社会的目的并不是为其他星球上的文明社会送去友谊，而是为了满足自己的野心。因此这架计算机成了地球的大独裁者，主宰了人类的命运。在科幻小说结尾的时候，主人公用一把利斧劈碎了这台魔鬼般的机器。

在我们看来，有趣的问题是，仙女座人究竟是如何控制地球的？仙女座人并不能始终直接控制计算机的所作所为，实际上他们也不可能知道计算机是否已经建造了，因为信息要等200光年以后才能到达仙女座。计算机完全能按它自己的主张独断专行。它甚至无法向它的主人征求基本决策的指令。所有的指令都必须预先制定输入，因为200光年的障碍是无法逾越的。从理论上讲，这台计算机的程序与对弈机应该是一样的，唯一的区别在于它更能吸收当地的信息和有更大的伸缩性。这是因为设计这样一台计算机不仅要在地球上使用，而且有可能要在任何

具有发达技术的星球上使用，每个星球上的情况千差万别，仙女座人是无从预先卜知的。

仙女座人必须利用在地球上建造计算机来作出每日的决策，同样道理，我们体内的基因必须造出大脑。但是，基因不仅像仙女座人那样发出编成电码的指令，它们自己也就是指令。它们无法直接操纵身体的原因是一样的，那就是时差。基因通过控制蛋白质合成来施加其影响。这是操纵世界十分有力的措施，但速度缓慢。为了建造一个胚胎，基因必须先耐着性子在蛋白质那里折腾好几个月。而身体各部位的活动则是速度极快。速度之快无法以月来计算，而只能以秒、毫秒来计算。世上的事瞬息万变，猫头鹰从头顶掠过，草丛窸窣作响暴露了猎物，霎那间神经系统作出了反应，随即肌肉收缩，或是保住了性命，或是丢掉了性命。基因的反应没有如此之迅速。它们和仙女座人一样，只能事先建造一台计算机，并根据“预先”所能想到的种种可能性，事先编入对应的程序。但是，就像对弈一样，生命变幻无穷，吉凶福祸难以预卜。因此，跟对弈程序设计者一样，基因给它们的生存机器“规定”的不是具体的知识，而只能是总战略和生存之道的窍门。

正如J·Z·杨指出的那样，基因不得不做一些类似预测的工作。任何一个生存机器当它还处在胚胎阶段时，就已潜存着未来生活中将要遇到的千难万险。谁能预料会有什么样的食肉动物正躲在灌木丛中，等着猎物，什么样的飞跑的猎物又会从面前跑过？人类和基因都无法预卜。尽管如此，有些基本预测还是能够作出的。北极熊的基因完全可以肯定它们尚未出生的生存机器将诞生在北极的冰天雪地里。它们的这种预测决不是先知，因为它们根本没有思维能力。它们只不过按照几千年的传统做法，为这些新生命准备一层厚厚的皮毛。这也正是它们自

已还存在于基因库里的原因。它们还预测了地面将是白雪皑皑，因此这层厚皮毛的颜色是白色的，能起很好的保护作用。假如北极圈内的气候突然发生了巨大的变化，一夜之间变成了酷热的热带沙漠，那么基因的预测就是错误的，它们因此将付出无法估算的代价。幼熊将会因炎热而死亡，体内的基因也会死亡。

*　　　*　　　*

一个十分有趣的预测前途的方法是模拟。如果一名将军想要知道他的军事计划是否最为合理，这就存在着一个预测问题。有着各种各样未知的因素，比如气候的变化、部队的士气以及敌人的对应措施等等。有一种检验一个计划优劣的方法，那就是演习一下。但是，一次军事行动会有许多个假想的计划，你不可能将每个计划都真刀真枪地演习一遍，因为人力和物力都会有一定的限制。所以，你只能进行模拟演习。在模拟演习中，可以假设一支“北军”与一支“南军”交战，士兵都使用空炮弹，不至于引起伤亡。即使是这样，演习也会很费时，而且耗资巨大。比较经济的做法是用锡制的士兵和玩具坦克在一张巨大的军用地图上布阵。

近年来，计算机在各种模拟试验中得到了广泛的应用，不仅在军事战略方面，而且在其他需要作未来预测的领域，如经济学、生态学、社会学和其他许多领域内也起着越来越重要的作用。模拟技术大致是这样一种情况。首先在计算机内设计一个模拟的某种环境模型。但这并不是说，你掀开盖子，会看见里面有一个缩小的与被模拟环境一模一样的模型。在对弈机内的存储库内并没有一幅“心理图画”，画中有一个摆着马、兵之类棋子的棋盘。实际上，棋盘和它的布局是由一些电子密码数字来表示的。对于我们来说，一幅地图就是世界的局部模型，被

压缩在二维空间之内。在计算机里，地图一般是由一系列的城市名和其他点来表示的，这些城市名和点都由两个表示经度和纬度的数字来代表。至于计算机如何在它的头脑里装下这个模型，并不重要，只要它能依照这个模型进行运算，进行操作，进行实验和用能为人所理解的语言向人如实汇报实验结果就可以了。通过模拟技术，可以决出战斗的胜负，可以使模拟的飞机飞行或坠毁，并可以预测经济政策的后果是振兴还是破坏。在每一次模拟中，计算机内部进行的整个计算过程相当于现实生活中整个行动过程的极小一部分。当然，模拟的模型有好的也有坏的，而好的模型也只是近似物。模拟绝不能准确无误地测定实际上将发生的情况，但是，一个好的模拟总比盲目尝试要好得多。模拟可以称之为代偿性试验，很可惜，这种说法早已为研究鼠类的心理学家抢先使用了。

既然模拟如此之好，那么，我们很自然地认为生存机器是它的第一个发现者。毕竟在人类出现之前，生存机器早已发明了人类工程中的许多技术，比如，焦距镜、抛物反射器、声波的频率分析，伺服控制、声纳和信息的缓冲存储以及无数其他有一长串名字的东西，这些技术细节与我们无关。那模拟又怎么样呢？这样说吧，当你对包含许多未知因素的未来事件踌躇不定时，你自然而然地就会求助某种形式的模拟。你在脑海中将每一种可能性都过滤一遍。你在自己的大脑里建起一个模型，当然这个模型并不包括世上万事万物，只包括了你认为与这件事有关的各种东西。你在想象中十分逼真地看到了这些东西，或是用抽象的方式来操纵它们。但不管是哪种情况，在你的大脑中并不存在一个你所想象的事件的实在的空间模型。但是，与计算机的情形一样，不管你用什么方式来表示世界模型都不重要。关键在于你是否能用这样一种方式来预测未来事件。能模拟未

来的生存机器，比只能在公开的尝试基础上学习的生存机器大大地进了一步。公开试验会造成时间和能量的浪费。公开失误常常带来毁灭性的后果。而模拟则要安全得多，也快得多了。

模拟能力不断进化的最高峰应该是主观意识。在我看来，主观意识为什么会发生是现代生物学面临的最深奥的谜。我们没有理由认为电子计算机能进行有意识的模拟，尽管将来会有这种可能性。当大脑对世界的模拟完备无缺到了包含一个自身的模型时，也许就产生了意识。很明显，生存机器的身体和各个关节也是它所模拟的世界的一个重要组成部份。我们还可以认为，由于同样原因，模拟本身也是所要模拟的世界的一部分。"自我感知"这个词也许可以用来描写这个过程，但我觉得它还不能令人满意地解释意识的进化过程，部分理由是它会导致恶性循环：如果有一个模型的模型，为什么就不可以有一个模型的模型的模型呢？……

不管意识引起什么样的哲学问题，就我们讨论的问题而言，意识可以被看作是一种进化趋势的高峰，这一趋势即把生存机器作为决定的执行者，从其最高主宰即基因那里解放出来。大脑不仅掌握了处理生存机器的日常事务的大权，而且还能预测未来，并能按照预测行动。它们甚至能够反抗基因的旨意，比方说，基因指示它们尽可能多生多育，但它们拒绝执行。不过，在这个方面，我们将会谈到人类是个很特别的样子。

所有这一切与利他主义和自私自利又有什么相干？我试图说明，动物的行为不管是利他的或是利己的，都是受到基因的间接的，但十分强大的控制。基因通过指示如何制造生存机器和它们的神经系统，对其行为进行最根本的控制。当然，每时每刻对下一步行动所作出的决定是由神经系统完成的。基因是最初的政策制定者，而大脑只是执行者。但是，随着大脑不断发

达，通过运用学习和模拟等技巧，它们获得了越来越多的实际决策权。尽管这种趋势的必然结果还没有在任何物种里发生，但是，我们可以断言，基因发给生存机器的是一个总体的政策指令。这个指令是：为了使我们活下去，你务必择善而行。

自私的米姆

在可及的宇宙里，物理定律被认为是普遍有效的。在生物学里，是否也有这种普遍适用的原则？当宇航员飞向遥远的星球寻找生命时，他们可望找到叫人难以想象的希奇古怪的生命。但是，不管是在哪儿发现的，也不管其化学基础是什么，究竟以什么作为衡量生命的标准呢？如果有这样的生命，它的化学成分不是以碳而是以硅，不是以水而是以氨的形式存在；如果有这样的生命，在负100摄氏度时就被烫死了；如果有这样的生命，它的基础不是化学成分而是电子反射环路，那么，试问是否还有什么衡量生命的标准吗？我实在不知道，要是非说一句不可的话，我将把赌注押在一个最基本的原理上：由于复制实体千方百计以图生存，生命才得以进化。基因也可以叫作DNA分子，它正是地球上无所不在的复制实体。也许还有其他的复制实体。如果真有的话，一旦满足某些条件，它们就会不可避免地成为一个进化过程的起点。

但是，我们是否非要去遥远的世界寻找其他复制者和其他进化模式？照我看，就在地球上，一种新的复制者近年来已经出现了。它正注视着我们，它还很幼小，在它的原始液里不灵活地飘浮着，然而，它的进化速度极快，令旧基因望尘莫及。

新的原始液就是人类文化之液。我们需要用一个名称来称

呼这种新的复制者，它要能表达文化遗传单位，或是模拟单位这样的概念。“mimeme”这个词源出自希腊语，但我想要一个单音节词，读上去有点象gene(基因)。如果我把“mimeme”缩略成“meme”，我希望研究古希腊语言的朋友们会原谅我。不管你会怎么说，这个新词在词形上多少还有些可取之处，它可以和“memory”(记忆)联系起来，与法语“même”也有点联系。这个词应和“cream”(奶油)押韵，读作“米姆”。

米姆的例子多不胜数，例如，曲调、思想、妙语、时装以及制陶和建筑工艺等。基因通过精子和卵子代代遗传，在基因库里不断繁殖。同样，米姆通过广义上说的模拟，从一个大脑传到另一个大脑，在米姆库内不断繁殖。如果一名科学家听说或谈到一个巧妙的构思，会马上将它传给同事和学生。他会在文章里和课堂上提到这个构思。一旦它被人们理解了，我们就可以说它得到了繁殖，在大脑之间广为传播。我的同事N·K·汉弗莱曾为本章写过一个初稿，他在初稿中十分简明地指出：“……从专门意义上说，应该认为米姆是有生命的结构，这不仅仅是比喻。如果你在我的大脑内种下一个米姆，这个米姆实际上就寄生在我的大脑之内，我的大脑成了米姆繁衍的容器，这同一个病毒寄生在基质细胞的遗传机能之内又有什么区别呢？这不是说说而已。例如，米姆作为世上个体神经系统中的一个结构已不知几百万次地从躯体方面实现了‘相信死后有生’的想法。”

* * *

我认为，互相适应的米姆复合体与互相适应的基因复合体是以同样的方式进行活动的。选择喜欢那些根据自己的优势利用文化环境的米姆，这个文化环境也包括了其他的米姆，这些米姆也要被选择。因此，米姆库逐渐成为一个进化过程稳定的

群体，新的米姆很难侵入。

我对米姆一直抱有一定的成见，然而，它们确实有可爱的一面。在我们死后，遗留下两样东西：基因和米姆。我们来到这个世上，实际上是当一辈子基因的生存机器，为基因传宗接代。但是，在三代之后，这个可悲的人生使命将会被渐渐淡忘。你的儿女，甚至你的孙儿女也许五官与你长得很像，也许同你一样酷爱音乐，也许头发的颜色也和你一样。但是，你的基因每隔一代就要减少一半，这样，随着一代一代的繁衍，用不了很久，你的基因的比例就会变得微乎其微。基因也许是不朽的，但是构成我们个人的基因结合体最终将会消亡。伊丽莎白二世是征服者威廉的嫡系后代，但是，也许她身上的基因没有一个是先王的。我们不应在繁殖中寻求不朽。

但是，如果你对世界文化有所贡献，提出过一个伟大的思想，谱过一首好曲子，发明过一个电花插头，或者作过一首好诗，它们就会永远流传下去，即使在你的基因已融入公库之后，它们仍旧永垂不朽。G·C·威廉曾经说过，苏格拉底恐怕早已没有什么基因遗留在当今世上了吧，可这又有什么关系呢？像苏格拉底、达·芬奇、哥白尼和马可尼这些伟人的米姆复合体仍然生机勃勃。

D·R·霍夫施塔特

反　　思

道金斯是图解还原论的大师，他以生动的笔触向我们描叙了生命和心灵如何由偶然产生的微生物经过获取资源的竞争

和无情的筛选，慢慢地在喧嚣的分子液中诞生的过程。还原论认为世界上的一切现象都可还原为物理学定律，因此，根本不可能存在所谓的"突生"特性。有的人为了说明这个特征，用了一个古老的但更加直言不讳的词"实现"，这就是说，更高层次的结构，是无法用支配其局部内容的定律来解释的。

请想象一下这样的情况：你的打字机(或者说你的洗衣机或复印机)坏了，你将它送到厂里去修理。一个月以后，工厂将打字机送回来了，他们已经把它拆开又装上了，丝毫没有差错。另外附了一张纸条，上面写道：对不起，所有的零件经检查都无毛病，但就是整座机器运转失灵。也许你会生气。如果机器不能工作，怎么可能每一个零件都是好的？肯定有地方出了毛病！常识告诉我们，差错出在日常生活的宏观领域里。

但是，如果倒过来从整体到局部，再从局部到局部的局部，这样一层又一层，那么，还原论的原则是否还适用呢？常识又会告诉我们：能够适用。然而，许多人却抱着这样的信念，比方说，"你不能由氢原子和氧原子的特征推断出水的特征"，"生物体远非其局部的总和"。人们经常将原子当作只有化学价的弹子球，没有对它作进一步的观察。事实证明，没有比这更错误的了。如果你认真观察微观粒子的特征，就会发现"物质"的数学特征会变得比以往更难以捕捉。让我们看一看理查德·马图克关于相互作用粒子的一段话：

> 关于多体问题讨论的一个比较合理的起点，也许应该是这样一个问题：在我们有问题之前，应需要多少个体？G·E·布朗教授曾经指出，对于那些对确定答案感兴趣的人来说，回顾一下历史，这个难题就迎刃而解了。在十八世纪的牛顿力学里，三体问题是无

法得到解答的。随着1910年广义相对论和1930年左右量子电动力学的诞生，双体和一体问题又陷入了困境。而在现代量子场理论中，对零体问题(真空)又找不出答案了。因此，如果我们执意要获得确定答案的话，那么，光是零体就够你受了，更不用谈多体了。

氧原子有8个电子。对这类原子的量子力学结构，我们是没有能力加以完全分析的。光是氢原子或氧原子的特征就十分复杂，无法形容，更不用说一个水分子的特征了，而氢原子和氧原子的这些特征正是水难以捉摸的特性的根源。我们可以用简化的原子模型，通过计算机模拟许多相互作用的分子，来研究这些特征。自然，原子模型做得越好，模拟就越真实。事实上，计算机在仅仅已知个别成分的特征情况下，可以通过模拟发现许多相同成分结合之后所产生的新特征，这已成为最常用的一种方法。在天体物理学领域内，计算机能够根据作为移动引力点的单个恒星的模型而产生对星系如何产生螺旋臂的新洞见。计算机也可以通过一个具有电磁相互作用结构的分子模型，研究固体、液体和气体如何振动、流动和改变状态的。

道金斯在其论文结尾时，就米姆问题提出了自己的看法，米姆是蛰居于心灵之中的软件复制者。接着，他又发展了这个观点，指出除基因外，米姆是生命延续的又一个媒介。但是，他没有注意到在中子星的表层，核粒子聚合、分离的速度比原子要快出成千上万倍。从理论上讲，核粒子的“化学成分”能够允许极其微小的自我复制结构存在，这些结构生命增殖速度极快，一眨眼就会增加许多许多倍，其复杂程度同地球上的微粒子一模一样。这种生命是否真的存在？假设它真的存在，我们是否能够发现它？这些问题的答案都无从知晓。但是，有人会

说，这也许好比一个小人国：世上仅数日，小人国内已千载！本书所选斯坦尼斯瓦夫·莱姆的几篇文章都持有这种看法。

我们向读者展现这样一个奇特的观点是为了提醒大家注意，在思考支持类似生命或思维活动的媒介的多样性时，务必不要存偏见。在下一篇对话中，我们又对这个问题进行了讨论，当然不像本篇这样异想天开。在这段对话中读者将看到意识从一个蚁群相互作用的不同层次上慢慢地出现了的情景。

11 前奏曲……
蚂蚁赋格曲

道格拉斯·R·霍夫施塔特

前奏曲……

阿基里斯和乌龟来到朋友蟹的住处，见到了蟹的朋友食蚁兽。相互寒暄一阵后，便坐下来喝茶。

乌龟：蟹先生，我们送您一件礼物。
蟹：太感谢了，你们来我这儿本不该带东西。
乌龟：只是表表我们的敬意。阿基里斯，请把礼物拿出来，行吗？
阿基里斯：当然可以。蟹先生，您好，我想您会喜欢它的。

（阿基里斯递给蟹一个精致的礼盒。方方的，很薄。蟹动手打开盒子。）

食蚁兽：不知是件什么礼物？

蟹：马上知道了(盒子打开了，蟹取出礼物)。两张唱片！太好了！怎么没有商标。噢，龟先生，这次又是您那“特殊”的礼物吗？

乌龟：您是说会毁坏唱机的唱片吗？这次不是。这唱片是定制的。全世界只有这一种了。以前从未有人听过。当然，巴赫演奏过除外。

蟹：巴赫什么时候演奏的？您到底是什么意思？

阿基里斯：蟹先生，要是龟先生告诉了您这唱片的来历，您会喜出望外的。

乌龟：阿基里斯，说下去，告诉他。

阿基里斯：可以吗？好家伙，那么我最好是查查笔记（拿出一张小卡片，清了清嗓子）。嗯哼，你们对数学领域卓越的新成果感兴趣吗？有了这些成果才有这些唱片。

蟹：有数学成果才有唱片？真奇怪！好吧，既然您已引起了我的兴趣，我一定要听听。

阿基里斯：好极了。那么。(停下呷了口茶，接着说)你们听说过臭名昭著的费马①“大定理”吗？

食蚁兽：我不敢肯定……它听来很熟悉，但我吃不准。

阿基里斯：这是个很简单的定理。皮埃尔德·费马是律师，但又是个业余数学家，当他在读一本丢番图②的经典著作《算术》时，看到这样一个方程：

---

① 费马(Fermat，1601—1665)，法国数学家，提出的一个非凡的定理也称作费马猜想，数学家们对这个命题历来感到困惑，因为他们既不能证明它成立，也不能证明它不成立。——译者

② 丢番图(Diophantus，246—330)，希腊亚历山大时期的数学家，以研究代数著名。他有个著名方程叫丢番图方程。此方程仅包括加法，乘法，取方程的系数为整数且其非零整数解的方程。——译者

$$a^2+b^2=c^2$$

他立即认识到，这个方程有无穷个解 a、b、c，并在一边写下了著名的旁注：

> 方程
>
> $$n^a+n^b=n^c$$
>
> 只有当$n=2$时，才得出a，b，c，n 正整数的解（而且，满足这个方程的三个一组的a，b，c，也有无穷个），但没有$n>2$的解。我已发现了一个确实非凡的证明，但遗憾的是，页边太小，写不下。

自从300年前的那天起，数学家们一直在徒劳地做着两件事：或者是证明费马的观点，从而维护他的声誉（虽然他享有很高的声誉，但有些怀疑论者认为，因为他从未真正证明过自己的观点而败坏了自己的名声）。或者是找出反例，驳斥费马的观点：用四个整数a、b、c和n，并且$n>2$来满足那个方程。直到最近，这两种尝试一次也没成功。诚然，这个定理已被证明出n的很多特定的值，尤其是，所有的n都达到125，000。

食蚁兽：如果它从未得到完满的证明，与其叫"定理"，还不如叫"猜想"。

阿基里斯：严格地说是这样。但按照传统习惯，人们一直是这么叫的。

蟹：有没有人最终解决了这个著名的问题？

阿基里斯：确实有。实际上乌龟先生就已经解决了。像以往一样，他施展了一下魔法就解决了。他不但发现了费马大定理的证法（从而不但证明这个名称是正确的，而且维护了他的

声誉），还找出了反例。这就表明，那些怀疑论者有着良好的直觉。

蟹：哎呀，太妙啦。这真是一个革命的发现。

食蚁兽：请别卖关子了。那些符合费马方程的不可思议的整数是什么？我对n的值特别感兴趣。

阿基里斯：噢，天哪，真难堪。你们相信吗？我把那些值写在家里那张巨大的纸上了。可惜那张纸太大了，不好拿。我真想把它带来让你们看看。如果对你们有帮助的话，我只记住一点，n的值是在π的连分数中不出现的唯一正整数。

蟹：噢，真遗憾，你没把那张纸带来。但我们没有理由怀疑您的话。

食蚁兽：可是，谁要看n的小数值来着？阿基里斯刚告诉我怎样找出n的小数值。啊，乌龟先生，鉴于您做出了具有划时代意义的发现，请接受我衷心的祝贺。

乌龟：谢谢。但我觉得比这结果本身更为重要的是，它应被立即应用。

蟹：我非常想听听怎么应用。因为我总是认为，数论是“数学皇后”——是纯而又纯的数学分支——是没法应用的数学分支。

乌龟：有这种看法的不止您一个人。但实际上，对于纯数学的一个分支甚或一些个别定理，何时以及如何在数学之外引起重大反响的问题，不可能作出囊括一切的说明，本例是说明这种现象的最好例子。

阿基里斯：乌龟先生模棱两可的结论在“音响再造”领域里是个突破。

食蚁兽：什么是“音响再造”？

阿基里斯：顾名思义，它将来自非常复杂的声源的音响信息进行再造。“音响再造”的典型工作是根据湖面上泛起的涟漪来再造石块落入湖水中产生的声音。

蟹：唷，这听起来不太可能。

阿基里斯：不能这么说。它实际上与人脑的活动非常相似。人脑根据耳鼓传递到耳蜗纤维的振动来再造另一个人声带所发出的声音。

蟹：我明白了。但我还不懂数论与本题又有什么相干，或者说这些与我的新唱片有何关系。

阿基里斯：在“音响再造”的数学里，有许多与丢番图方程多种解法有关的问题。乌龟先生几年来一直想从目前大气层所有分子运动的计算中找出一种再造巴赫200多年前演奏大键琴声音的方法。

食蚁兽：那当然是不可能的。那些声音已无法挽回地消失了，永远地消失了。

阿基里斯：天真的人就是这样说的……然而乌龟先生多年来一直致力于此问题的研究，并终于发现这个问题与下列方程的多种解有关：

$$a^n+b^n=c^n$$

解为正整数，且$n>2$。

乌龟：当然，我可以解释这个方程是怎样得出来的。但您听了肯定会厌烦。

阿基里斯：结果是，音响再造理论预示，巴赫的声音可以从空气中所有分子运动中再造出来，条件是要么这个方程至少有一个解。

蟹：真神奇！

食蚁兽：真怪！

乌龟：真想不到！

阿基里斯：我是想说"条件是要么这个方程至少有一个解，要么证明这个方程无解！"因而，乌龟先生谨慎地着手从两方面同时解决这个问题。事实证明，找出反例是找出证明所必不可少的一步。这样，就可以直接做到由此及彼。

蟹：那怎么会呢？

乌龟：我已向你们说明了费马大定理的证明如果存在的话，其结构能用一个漂亮的公式来表示。这个公式，事实上取决于某个方程解的各个值。当我得出这第二个方程时，惊讶地发现它原来就是费马方程。这是一个形式和内容之间有趣的偶然联系。因此，当我找出反例时，只须用这些数字作为我证明这个方程无解的根据。这么一想，您就觉得相当简单。我想象不出为什么以前从未有人发现这个结论。

阿基里斯：由于这个意想不到的珍贵数学成就。乌龟先生就能使他长期以来梦寐以求的"音响再造"付诸实施。给蟹先生的这个礼物清楚表明，这些抽象工作已经实现了。

蟹：它总不会是巴赫演奏大键琴的录音唱片吧！

阿基里斯：对不起，正是这么一回事。这是一套两张的唱片，是塞巴斯蒂安·巴赫的《平均律钢琴曲集》全集的上、下两卷。每张唱片录有24支前奏曲和赋格曲——每支曲子都有一个大调，一个小调。

蟹：啊！我们得马上把这贵重的唱片放放听听。我该怎么感谢你俩呢？

乌龟：您已经用这香茗盛情款待了我们。

（蟹拙出一张唱片，开始在唱机上播放起来，顿时

一位熟练的大键琴家演奏的那难以置信的美妙乐曲在室内回荡着。真难以想象，非常逼真。人们甚至听见——是听见，还是想象？——巴赫一边演奏一边轻声和唱……）

蟹：你们谁边听边看乐谱？我恰好有一个《平均律钢琴曲集》的珍本，我的一位老师特意为它设计了饰字，他是一个了不起的书法家。

乌龟：我很想欣赏这乐谱。

（蟹走到他那精致的木制书柜前，打开玻璃门，拿出两本厚书。）

蟹：给您，乌龟先生。我对这书中的精美图案向来是外行，也许您的礼物会使我成为内行。

乌龟：我希望如此。

食蚁兽：您注意到了吗，这些作品中的前奏曲总是为紧接着的赋格曲完美地定了调子？

蟹：注意到了。虽然难于用文字表达，但它们之间确实有种微妙的关系。即使前奏曲和赋格曲没有共同的主题调，然而在它们之间的确有些不可捉摸的抽象特征，把它们紧紧地连在一起。

乌龟：在前奏曲和赋格曲之间有片刻戏剧性的停顿——赋格主题就在这停顿中马上以单音奏出，接着便融入越来越响的、神奇的优美和声之中。

阿基里斯：我明白您的意思了。我虽然没有听过很多前奏曲和赋格曲，但我觉得短暂的寂静激动人心。现在我可以来猜

测老巴赫的意图了。比如：我总在想，赋格曲的速度是快板还是慢板？是6/8拍还是4/4拍？是 3 个声部，还是 5 个声部，还是 4 个声部？第一个声部开始了……这是个多么令人陶醉的时刻。

蟹：是这样，早已逝去的青春岁月我还记得清清楚楚。那时，每首新的前奏曲和赋格曲都会使我心情激荡，它们所包含的新奇、优美及许多意想不到的变化使我激动万分。

阿基里斯：那现在呢？那时的激动全部消逝了吗？

蟹：是的。听多了也就觉得无所谓了，一切激动人心的东西到后来都是这样。但这种熟悉又有一种深度，自会产生某些补偿。例如，我总会发觉一些以前所未注意到的新奇之处。

阿基里斯：您忽略的是主题吗？

蟹：可能是。特别是当它转换成和隐藏在其他声部中之后，又不知从何处突然冒了出来。也有些令人惊讶的变奏，使人百听不厌。真不知道老巴赫是怎么构思出这么美妙的变奏来的。

阿基里斯：我初次听《平均律钢琴曲集》时真为之陶醉，现在您说听这乐曲仍有所得益，我感到很高兴——尽管我悲叹这样的阶段并不能持久。

蟹：噢，您也不必担心这陶醉会永远消逝。有一件事值得我们感到欣慰，就在您认为青年时代的陶醉已逝去的时候，它却每每又复萌了。它正是受到外来因素的激发。

阿基里斯：真的吗？此话怎讲？

蟹：比如，通过某个人的耳朵听见它，而那个人又是第一次听——如像您，阿基里斯一样。不知怎的，那激动传播开来，我便能再次感到激动。

阿基里斯：真有趣。这种激动在您身上处于潜伏的状态，但是单靠您自己是无法把它从您的潜意识中引出来的。

蟹：正是如此。重新体验那种激动的潜力以某种不为人所知的方式在我的大脑结构中"编了码"，但我无法随意地将它唤出，必须等待偶然的机遇来激发那种能力。

阿基里斯：我有个关于赋格曲的问题，提它不太好意思。我一直在想，由于我刚开始听赋格曲，你们经常听赋格曲的人能否在此方面给我些帮助呢？……

乌龟：尽管我懂得也不多，但如果对您有帮助的话，我当然愿意效劳。

阿基里斯：啊，谢谢。让我从某一个角度谈这问题。你知道M·L·埃舍尔画的名叫《魔带立方架》那张画吗？

乌龟：您是要说在那个立方块的外面绕了几圈带子，带子上面有一个个圆泡泡。您刚觉得这些圆泡是凸起的时，它们又显得是凹下去的，反之也如此，对不对？

阿基里斯：一点不错。

蟹：我记得那张画。那些圆泡似乎总是在凹面和凸面间来回变动，是凹是凸，根据您观察的角度而定。不可能看到它们既有凸又有凹——不知怎的，人的大脑不允许这样。人们只可以在两个互为排斥的"方式"中选择一个。

阿基里斯：正是这样，听赋格曲时我似乎也发现了两个类似的方式。这两个方式是：每次或者是听一个单声部，或者是听所有声部的整个的效果，不去区分其中单个的声部。我已试过这两种方式了，但令人沮丧的是，每种方式都把另一种掩盖了。我简直没法做到既听单声部又听整个效果。我发现，我多少是同时又无意识地在一个方式和另一个方式间来回移动。

食蚁兽：就像您观察魔带一样，是吗？

阿基里斯：是这样。我刚才在纳闷……我对听赋格曲的两种方式的描述是否清楚地表明我自己是这样一个天真、没有经验的听众，甚至没能把握住更深层的感知方式，因为这种方式是我力所不能及的。

乌龟：不，完全不是，阿基里斯，我只是说我自己。我发觉自己也在一个方式和另一个方式之间来回变换，应以哪个方式为主我也无法自觉控制。不知在座的其他各位是否也有类似的体验。

蟹：肯定有。这是个逗人的现象，因为您感觉到赋格曲的本质在您身边萦回，而您又不能完全理解，因为您不可能同时在两个方面作出反应。

食蚁兽：赋格曲有这样的有趣特性，它们每个声部本身就是一首乐曲，因此不妨把赋格曲看作是几首不同乐曲的组合，它们都基于一个主题而同时奏出。那就该由听者（或他的潜意识）来判定它应被理解为一个单位，还是独立而又协调的部分的组合。

阿基里斯：您说那些部分是"独立的"，其实不可能。它们之间必须互为协调，否则，合奏起来就会产生乱七八糟的声音——实际上绝不是这么一回事。

食蚁兽：换个说法也许更明白一些：如果您单独听每个声部时，您会发觉它似乎可以自成一体。它可以独立成章，我说的"独立"就是这意思。不过，您说得也对，您说每个单独的有意义的声部与其他的声部以极度有序的方式相融合，就能产生出优美的浑然一体的乐曲。写出优美的赋格曲的艺术恰恰靠这种能力，创造几个不同的声部，每个声部似乎是自成一章，非常完美，但合奏起来又浑然一体，绝没有

牵强之感。再说，把一首赋格曲作为整体或作为部分声部听的二分法是那种适用于许多由低级层次构造起来的结构的极普遍二分法的一个特例。

阿基里斯：真的吗？您是指我说的两个"方式"具有更加普遍的应用性，并不局限于听赋格曲吗？

食蚁兽：完全正确。

阿基里斯：我不明白这怎么会呢？我认为这跟我们时而把一件事物从整体上来观察，时而又从部分的集合上来观察密切相关。但我只是在听赋格曲方面才碰到这种现象。

乌龟：啾，天哪！您瞧，我正好翻到现在奏的那页，看到了赋格曲开始的那页上有幅优美的插图。

蟹：我以前从未看过这插图。为什么不递过来让我们也瞧瞧？

（乌龟把书传给大家，四个读者中每位都以一种独特的方式看它——这个从远处看，那个从近处瞧，每位都这样或那样地摇着头，迷惑不解。最后书又传回到乌龟手里，他颇为仔细地审视着。）

阿基里斯：对了，我想前奏曲快完了。我在想，当我听这首赋格曲时，不知会不会对"听赋格曲的正确方法是什么？是把它作为整体还是部分的集合？"的问题有更深的理解。

乌龟：只要仔细听，您一定能办到。

（前奏曲完了，大家沉默不语，接着……

［继续演奏］

……蚂蚁赋格曲

……接着，赋格曲的四个声部一个接一个地奏起来了。)

阿基里斯：我知道，你们几个不会相信这点，但问题的答案就在面前，它隐藏在这张图画中，它只是个简单的词——但又是个多么重要的词啊！它是"MU"！

蟹：我知道你们几个人不会相信这点，但问题的答案就在面前，它隐藏在这张图画中，它只是个简单的词——但又是个多么重要的词啊！它是"HOLISM"(整体论)！

阿基里斯：等等，您一定看错了，这张图画明明白白，那上面的字是"MU"而不是"HOLISM"！

蟹：对不起，我的视力极好。请您再看看，然后告诉我，这张画上的字与我说的是一样还是不一样。

食蚁兽：我知道，你们几个人不会相信这点，但问题的答案就在面前，它隐藏在这张图画中，它只是个简单的词——但又是个多么重要的词啊！它是"REDUCTIONISM"(还原论)！

蟹：您等等。您一定看错了。这张画上面明明写着"HOLISM"不是"REDUCTIONISM"！

阿基里斯：你们看错了。既不是"HOLISM"也不是"REDUCTIONISM"而是"MU"，肯定是。

食蚁兽：对不起，我的视力极好，请再看看，告诉我这张画与我说的是一样还是不一样。

阿基里斯：难道你们没看见这张画是由两部分组成，每一部分各是一个字母？

蟹：您说有两部分这对了，但它们各是什么，您却说错了。左

边的那部分完全是由三个重复的词"HOLISM"组成；而右边那部分是由许多重复的词组成，字体非常小。我不知道为什么在这两部分中字母大小各不相同，但我看见的是明明白白的，是"HOLISM"，非常清楚。你怎么看见别的东西，这我就不明白了。

食蚁兽：您说分两部分是对的，但它们各是什么，您却说错了。左边部分完全是由许多重复的"REDUCTIONISM"组成，而右边部分是由一个"REDUCTIONISM"组成，字体很大，我不知道，为什么在这两部分中字母大小各不相同，但我看见的是明明白白的，是"REDUCTIONISM"。它非常清楚。您怎么还看到别的东西，这我就不明白了。

阿基里斯：我知道，这是怎么回事。你们每位看到的字母都构成了其他字母，或由其他字母所组成。在左边的部分里，的确有三个"HOLISM"。它们每个又由许多更小的词"REDUCTIONISM"组成。作为补充的形式，在右边部分中，确实有一个"REDUCTIONISM"。但它是由小写的许多"HOLISM"组成。这就是绝妙之处。而你们俩愚蠢的争吵，实际上都是只见树木不见森林的结果。你们瞧，当理解的正确方法超越了是"HOLISM"还是"REDUCTIONISM"这个问题本身，而回答"MU"时，你们的争论还有什么意义呢？

蟹：阿基里斯，现在按您的描述我看明白这张图了。但我弄不懂您说的"超越问题"是什么意思？

食蚁兽：阿基里斯，现在按您的描述我看明白这张图了。但我仍不懂您那古怪的词"MU"是什么意思？

阿基里斯：如果你们首先答应我的要求，告诉我holism和reductionism这两个古怪的词是什么意思，我会很乐意回答你

们的。

蟹：整体论是世上最容易把握的思想。它不过是认为，“整体大于部分之和”。凡是理智健全的人都不会否认“整体论”。

食蚁兽：还原论是世上最容易把握的思想。它不过是认为“如果您理解了整体的部分和它们之和的本质，也就完全理解了这个整体”。就是蠢人也不会拒绝这个观点。

蟹：我反对还原论。比如，请您告诉我怎样用还原论的观点来理解大脑。任何有关大脑的还原论解释最终将无法说明大脑的意识从何而来。

食蚁兽：我不赞成整体论。比如，我请您说说，对蚁群的整体描述怎么会比对蚁群的成员及其作用和相互间的关系的描述更符合蚁群的实情。任何整体论的蚁群解释最终都无法说明蚁群意识的来源。

阿基里斯：噢，不。我最不愿意挑起另一场争论了。不管怎么说，既然我了解了争端，我相信，我对“MU”的解释将大大有助于解决争端。你们知道吗，“MU”是古代禅宗的回答法。即是，提出问题又不回答问题。这儿问题好像是：“该用整体论，还是该用还原论来理解世界？”这儿“MU”的回答否认了这个问题的前提，认为两种理论都不能用。通过不回答问题，它揭示了一个更广泛的真理：有一个既适合于整体论也适合于还原论解释的更广阔背景。

食蚁兽：这太荒唐了！您的“MU”就同母牛叫一样蠢。我根本不要听禅宗这空洞无聊的话。

蟹：大可笑了！您的“MU”就像猫的叫声一样蠢。我根本不要听禅宗这空洞无聊的话。

阿基里斯：天哪！我们现在竟毫无进展。乌龟先生，你为什么不吱声？这使我太不自在了。我想您一定能解开这难题的，

是吗？

乌龟：我知道，你们几位不会相信这点。但问题的答案就在面前，就在这张图里，它只是个简单的词，而且是个挺重要的词"MU"！

（说到这儿，那支赋格曲的第四个声部奏响了，一开始就是个高八度音。）

阿基里斯：噢，乌龟先生，这次您真令我失望，我总觉得您看事一贯洞察秋毫，所以能解决这个困境，但看来您不过和我一样，当然就这一次。能和您看法相同，我很高兴。

乌龟：对不起，我没听清，但我的眼力极好，请再看看，然后告诉我是否这张图和我说的不一样。

阿基里斯：当然一样！您只不过重复了我最初的观察结果。

乌龟：阿基里斯，也许"MU"在这画中的层次比您想象得更深——（打个比喻）低一个音阶。但现在我想，是不是可以用抽象方法解决这场争论。我想看到整体论和还原论的观点能更明确地展开些。这样，就会为决断的依据提供更多的材料。例如，我非常想听听有关蚁群的还原论描述。

蟹：也许食蚁兽博士会告诉您他对这方面的看法的。他毕竟是这方面的专家。

乌龟：博士，我相信我们能从您这样的蚁学专家这里学到不少东西，您能用还原论的观点给我们再多讲一些有关蚁群的事吗？

食蚁兽：当然可以。正像蟹先生刚才说的那样，我的职业使我较深入地研究了蚁群。

阿基里斯：我可以想象得出！食蚁兽的职业好像与蚁群专家的

工作是一致的。

食蚁兽：对不起，你说什么？“食蚁兽”不是我的职业，而是我的同类。就职业来说，我是个蚁群外科医生。我专门用外科切除手术来治疗蚁群的神经错乱。

阿基里斯：噢，我明白了。那么您说的蚁群“神经错乱”是什么意思？

食蚁兽：我的大多数病人都患有一种语言障碍症。您知道吗，蚁群必须在日常环境下搜寻词汇。这也许是相当悲惨的。我想改善这种情况。嗯，通过切除蚁群中有缺陷的部分。这种手术有时非常复杂，事先当然还需要作几年的研究。

阿基里斯：但是您能肯定在患语言障碍症前一定有说话的官能吗？

食蚁兽：对。

阿基里斯：因为蚁群本不具备这种官能，所以我现在有点糊涂了。

蟹：太糟糕了。阿基里斯，上星期食蚁兽博士和希拉里大婶来我家作客，那天该请您作陪就好了。

阿基里斯：蟹先生，希拉里是您的大婶吗？

蟹：噢，不是，其实她不是谁的大婶。

食蚁兽：但这可怜的乖乖总让人这么称呼她，甚至让陌生人也这么叫。这是她许多讨人喜欢的怪癖中的一个。

蟹：是的。希拉里大婶非常古怪，不过很讨人喜欢。真遗憾，我上星期没把您请来见见她。

食蚁兽：她是我有幸结识的受过最良好教育的蚁群之一。我们俩就各种话题谈到很晚。

阿基里斯：我还以为食蚁兽不是蚂蚁智力活动的保护者，而是专门吞吃蚂蚁的呢！

食蚁兽：当然这两者是不矛盾的。我与蚁群交情很好。我只吃蚂蚁，不吃蚁群。这对我和蚁群双方都有益。

阿基里斯：那怎么可能——

乌龟：怎么可能——

阿基里斯：吃掉蚂蚁能对蚁群有利？

蟹：那怎么可能——

乌龟：烧了森林能对森林有利吗？

食蚁兽：那怎么可能——

蟹：剪掉树枝对树有利吗？

食蚁兽：理发对阿基里斯有利吗？

乌龟：你们几个忙于争论，没注意到巴赫的赋格曲中刚才奏出的那段的提前进入。

阿基里斯：什么是提前进入？

乌龟：噢，对不起。我还以为您知道这个词。它是指一个主题接连进入一个又一个声部，中间有一小会间歇。

阿基里斯：如果我多听赋格曲，很快也会知道这些的，会自己分辨出它们，不用别人指教。

乌龟：对不起，朋友们。很抱歉打断了你们的谈话。食蚁兽博士正要解释吃蚂蚁和与蚁群为友完全是一回事。

阿基里斯：对了，我现在还不太明白。您每天吃一定数量的蚂蚁与提高蚁群的健康水平有什么关系？我更不解的是，您怎么可能与蚁群对话。这绝不可能。一个蚁群就是一群到处乱爬的，寻找食物和筑巢的单个的蚂蚁。

食蚁兽：阿基里斯，如果您要只见树木不见森林的话，也可以这样说。实际上，作为整体的蚁群是个界限明确的单位，有自己的特性，有时还有自己的语言。

阿基里斯：我发现很难想象出，如果我在森林中大声呼喊，能

听到蚁群回答我的声音。

食蚁兽：蠢家伙！不是那么回事。蚁群不靠大声说话而靠书写来交谈，您知道蚂蚁怎样在身后留下忽东忽西的印痕的吗？

阿基里斯：哦，知道，蚂蚁通常穿过厨房的下水道再爬进到我的桃酱里。

食蚁兽：事实上，有些印痕包含着密码式的信息。如果您知道这个语言体系，就能像读书一样理解它们。

阿基里斯：真了不起。您能把信息反馈给他们吗？

食蚁兽：没问题。正是掌握了这一点，我才能和希拉里大婶交谈几小时。我用一根枝条在潮湿的地上划下痕迹，看着蚂蚁顺这痕迹爬行。现在，一条新痕迹开始形成。我高兴地看着它继续延伸下去。并且我同时还在预见它将怎样延伸（我的预见多数是错的）。当痕迹结束时，我就知道希拉里大婶的想法了，然后该轮到我回答了。

阿基里斯：我想，在这群蚂蚁里一定有几个极聪明伶俐的家伙。

食蚁兽：我觉得您还不太辨得清层次的差别。就像您决不能把单棵的树和森林混为一谈一样，现在也决不能把一个蚂蚁当作整个蚁群。要知道，希拉里大婶手下的所有蚂蚁都极笨。它们没法用交谈来解救它们的小命！

阿基里斯：那么，交谈的能力从何而来呢？它一定在蚁群内的某个地方！我不明白，如果希拉里大婶能用机智的笑话与您交谈几小时，那这些蚂蚁怎么能这样蠢笨呢。

乌龟：在我看来，这和大脑由神经原组成的情形差不多。要解释人能进行有智力的对话，肯定没人会认为非得每个大脑细胞都须成为有智力的人不可。

阿基里斯：不，显然不是这样。就脑细胞而论，我完全明白您的观点。而蚂蚁就另当别论了。我的意思是，蚂蚁是任意地、完全随机地到处乱爬，有时遇到一小块食物。它们想怎么爬，就怎么爬。从总体上，我根本看不出它们的行为有什么连贯而有序之处。尤其找不到像交谈所必需的大脑活动那样连贯有序的特点。

蟹：我似乎觉得，蚂蚁只是在一定限制内是自由的。例如：它们自由地爬行，互相碰撞，衔些小东西，留下一些尾痕等等。但它们从未迈出过那个小世界——它们所在的那个蚂蚁系统一步。它们从未想到过要这样做。因为它们没有想象这类事的能力。因此，蚂蚁是非常可靠的成员。你可以依靠它们用某种方式来完成某种任务。

阿基里斯：但即使如此，在那些限度内，它们仍是自由的。只是随意地爬行。它们从未想到过食蚁兽博士断言它们只是其成员的那个高级存在者的思维机制。

食蚁兽：噢，阿基里斯，有一点你没认识到，那就是统计资料的规律性。

阿基里斯：那是怎么回事？

食蚁兽：例如，尽管作为个体的蚂蚁好像以随机的方式爬行，然而它们有总的趋向，大批蚂蚁爬行的总趋向可以从那混沌状态中看出来。

阿基里斯：嗯，我明白您的意思了。事实上，蚂蚁的痕迹是说明这种现象的最好例子。您的确无法预见单个蚂蚁的爬行方向，然而痕迹本身好像是明确而稳定的。当然这必定意味着个别的蚂蚁不完全是随便乱爬的。

食蚁兽：太对了，阿基里斯。蚂蚁之间进行的某种程度的交流正好能阻止它们完全漫无目的地乱爬。通过这种最低级的

交流，它们可以彼此提醒，自己并不是孤独的，而是同伙伴们合作的。需要有许多蚂蚁相互间不断加强这种交流，才能从事长时间或短时间的活动，例如制造痕迹就是如此。我对人脑活动的极模糊理解使我相信，这与神经原的激活有点类似。蟹先生，要使一个神经原被激活，不是先得使另外一组神经原被激活吗？

蟹：当然是的。就拿阿基里斯大脑中的神经原来说吧。每个神经原从附着于它的输入神经上的那些神经原接受信号，当输入的总量超过临界限度时，那个神经原就会激活，将自己的输出脉冲传送给其他的神经原。于是这些神经原也激活了，沿着那根神经一直将信号传送下去。神经反射按照阿基里斯的方式无情而迅猛地发生，势如饿蚊铺天盖地，而且更不可思议，在感觉信号输入之前，每一转，每一动都由阿基里斯脑中的神经结构预定好了。

阿基里斯：平时我认为我能控制我的思维。可是按您刚才的那种说法，听起来"我"好像只是这些神经结构和自然法则的产物。我的自我充其量不过是个由自然法则控制的有机体的副产品，最糟不过是个由我的曲解认识所产生的人为概念。换句话说，您使我感到不知自己是谁，是什么，如果我确实存在的话。

乌龟：随着讨论的深入，您会更明白的。但是，食蚁兽博士，您对这个比喻怎么理解？

食蚁兽：我以前知道，在这两个极不同的系统里存在着类似的事物。现在我更清楚了。只有当具备一定数量的蚂蚁时，才会出现那种有组织的群体现象，例如制造痕迹。如果最初是在某个场所由几只蚂蚁也许是随意地爬出了痕迹，那么后来可能出现两种情况：一种是，爬了一会，就各自爬散

了。痕迹终未形成。

阿基里斯：是不是蚂蚁不够多，没法进行下去？

食蚁兽：对。可能出现的另一种情况是，来了许许多多的蚂蚁，事情就像滚雪球般地进行下去，越来越多的蚂蚁卷了进来。在这种情况下，形成了一支"队伍"。所有成员都在贯彻一个单一的计划，这计划可能是制造痕迹，也可能是采集食物，也许是保护巢穴。尽管这种工程规模很小时显得非常简单，但要是大规模进行就会产生极复杂的结果。

阿基里斯：从您的概述中，我可以大致把握从混沌到有序的总概念，但这与交谈能力还相去甚远。当气体分子随机地相互弹撞时，有序归根结底也是从无序中产生的，不过所产生的是一种不定形的，但以下述三种参数为特征的物体：体积、压力和温度。但这距了解宇宙或者谈论宇宙的能力还很遥远。

食蚁兽：这更突出了对蚁群活动的解释与对容器中气体活动的解释之间的一个有趣差别。人们只须通过计算分子运动的统计性能就能解释气体活动。我们没有必要讨论比分子更高的元素结构。另一方面，只有当您深入到几层结构之后，才能理解蚁群的活动。

阿基里斯：我明白您的意思了。在气体中，您一下子就能从最低的层次——分子；认识到最高的层次——丰富的气体。没有什么中介组织层次。在蚁群中怎么会有中介层次的组织活动呢？

食蚁兽：这与蚁群中有不同种蚂蚁的存在有关。

阿基里斯：是的，我想，我听说过这些，它们叫"种姓"，是吗？

食蚁兽：对。除了蚁后外，还有些公蚁，它们是不管蚁巢维修这类事的，然而……

阿基里斯：当然还有一些兵蚁——那些反对共产主义的光荣战士们！

蟹：晦……，我认为不对，阿基里斯，蚁群内部是完全共产的，为什么这些兵蚁要反对共产主义呢？食蚁兽博士，我说的对吗？

食蚁兽：对。蟹先生，有关蚁群的看法您是对的。它们的确有点儿以共产主义原则为基础。但阿基里斯关于兵蚁的说法未免有点天真了。其实，所谓"兵蚁"根本就不擅长打仗。这种蚂蚁长着大脑袋，行动笨拙迟钝，用很厉害的嘴咬东西，谈不上有什么光荣。在一个真正共产主义的国度里，配称光荣的是那些工蚁。是它们承担了大部分日常琐碎的工作·诸如采食、狩猎、养育幼蚁。甚至还承担了大部分作战的义务。

阿基里斯：呸，这真荒唐，兵蚁不会打仗！

食蚁兽：就像我刚才说的，他们其实根本不是兵蚁，工蚁才是兵蚁，兵蚁只是些懒惰的大脑袋。

阿基里斯：太丢人了！如果我是个蚂蚁，我就给它们的军队制订纪律，我要严惩这些大脑袋。

乌龟：如果您是个蚂蚁？您这样顺从的部下怎么可能是蚂蚁？根本没法把你的大脑绘入蚂蚁大脑里，所以我看考虑这个问题毫无成效。更合理的是把您的大脑绘入一个蚁群内……但我们别扯远了。让食蚁兽博士继续讲他那最富启发的种姓及其在高层组织中作用的理论吧！

食蚁兽：好吧。一个蚁群中有多种任务需要完成。个体的蚂蚁逐渐特化。通常一个蚂蚁的特化是随着蚁龄而改变的。当然这也取决于蚂蚁的种姓。在任何时候，在蚁群的任何一小块区域内，蚂蚁都是种类俱全的。当然一个种姓在有些

地方很稀少，在另一地方却很密集。

蟹：特定种姓的密布或特化是偶然的吗？某种蚂蚁在某处较集中而在别处则较稀少，这是什么原因？

食蚁兽：您提出这个问题我很高兴。这对了解蚁群如何思维有决定意义。事实上，在蚁群内需要很长时间才演化出非常微妙的种姓分布状况。正是这种分布使蚁群获得了作为与我交谈能力之基础的那种复杂特性。

阿基里斯：我觉得蚂蚁来回不断地爬会完全破坏微妙分布的可能性。蚂蚁的随机爬动会立刻把任何微妙的分布都打乱，正像气体中分子的微妙构成不一会儿就被来自四面八方的任意碰撞破坏了一样。

食蚁兽：在一个蚁群里，情况完全相反。事实上，正是蚁群中众多蚂蚁的来回爬动使种姓分布适应于变化着的环境，从而保持了微妙的种姓分布。要知道，种姓分布不能老是一个固定的模式，它必须以不断变化的形式从某种程度上反映该蚁群所面临的现状。正是蚁群内部的运动更新了种姓分布状况，以适应蚁群所面临的现状。

乌龟：您能举个例子吗？

食蚁兽：可以。当我作为一个食蚁兽去拜访希拉里大婶时，所有愚蠢的蚂蚁一闻到我的气味，都显得惊慌失措。当然这意味着，它们开始到处乱爬，把我来以前的分布状况都打乱了。

阿基里斯：这是可以理解的，因为您是蚁群的可怕天敌。

食蚁兽：噢，不。我必须重申，我根本不是蚁群的天敌，而是希拉里大婶的好朋友。我也很喜欢她。就算您说得对，蚁群中所有个别蚂蚁都怕我，但这完全是另一回事。总之，您瞧，由于我的到来，引起了蚂蚁们的慌乱，完全改变了蚂

蚁内部的分布情况。

阿基里斯：这就清楚了。

食蚁兽：这就是我所说的更新分布状况。新的分布反映了我的存在。可以把这种从旧状态到新状态的转变说成是给蚁群增加"一点知识"。

阿基里斯：您怎么能把蚁群内部不同蚁种的分布说成是"一点知识"呢？

食蚁兽：这是很关键的一点，需要作详尽阐述。您看，接下来是您打算怎样描述这种种姓分布。假如您仍然从低层次（个体蚂蚁）方面考虑，那就会只见树木不见森林。这只是个极度微观的层次。当您微观地考虑问题时，就肯定会忽略一些宏观特征。您必须寻找恰当的高层次框架，以便用它来描述这种种姓分布。只有这样，才能明白种姓分布怎能译解成许多知识。

阿基里斯：那么，您怎样寻找适当规模的单位来描述蚁群的现状呢？

食蚁兽：好吧，让我们从头讲起。当蚂蚁需要做某件事时，就组成一些小小的"队伍"，共同完成一些短暂的任务。就像我早已说过的那样，小群蚂蚁不断地形成和解散。那些存在较长时间的就是那些队伍。它们所以不解散是因为的确有事要做。

阿基里斯：您以前说过，如果蚂蚁的数量超过了一定的限度，就会形成蚁群，而现在您又说，它们必须有事要做，才会形成蚁群。

食蚁兽：这两种说法是一样的。举个例子，在采集食物时，如果某处有点微不足道的食物被几只四处爬的蚂蚁发现了，它们就会奔走相告，响应而来的蚂蚁数量与食物的大小是

成比例的。因为，一点点食物不会吸引过多的蚂蚁。这正是我说的没什么事可做的意思。因为食物太少，得不到重视。

阿基里斯：我明白了。这就是说那些"蚁队"所属的层次介于单个蚂蚁层次与蚁群层次之间。

食蚁兽：对，有一种特殊的蚁队，我称之为"信号蚁队"。更高层的结构都以信号蚁队为基础。事实上，所有更高的实体都是动作谐调的信号蚁队汇集。构成更高层蚁队的成员不是蚂蚁，而是低层次的蚁队。最低层次的蚁队就是信号蚁队，信号以下则是个别的蚂蚁。

阿基里斯：信号蚁队为什么会得到这个暗示性的名称？

食蚁兽：这取决于它们的作用。信号蚁队的作用是把不同种类的蚂蚁传送到蚁群的适当部位。所以对信号蚁队的典型描述是这样的：由于超出生存所需的限度而产生了信号蚁队，然后，它们迁移到蚁群的远处某地，或多或少解散一些自己的个体成员，任其独自行动。

阿基里斯：这就像波涛从远处带来沙石和海草，又把它们抛撒在高处干燥的海岸上。

食蚁兽：是有点像。因为蚁队中确能沉淀下一些来自远处的东西。但是，波涛中的海水还能再回到海中，而一个信号蚁队所携带的物质则不能回去，因为它是由蚂蚁本身构成的。

乌龟：我认为，信号蚁队正是在蚁群中最需要这种蚂蚁的地方丧失了它的内聚力。

食蚁兽：那当然了！

阿基里斯：当然！我认为不清楚的是，信号蚁队怎会总是应该去需要它的地方。即使它走的方向正确，它又怎么知道在

哪儿解散？怎么知道自己已到达目的地了呢？

食蚁兽：这些是非常重要的问题。因为它们牵涉到解释信号蚁队的有目的活动或似乎是有目的的活动，根据描述，人们很容易认为，信号行为的特征是满足一种需要，并称这种活动是“有目的的”。但是您也可以不这样看。

阿基里斯：噢，等一等。活动不是有目的的，就是无目的的，我不明白您怎么能两者兼得。

食蚁兽：让我来解释一下我的观察事物的方法，然后看您是否赞同。信号蚁队一旦形成，就不会意识到该在哪个特定的方向转向。但这时微妙的种姓分布发挥了决定性的作用。是它决定了信号蚁队在蚁群中的运动，以及该保持稳定多久和在哪儿解散。

阿基里斯：所以说，一切都取决于种姓分布，是吗？

食蚁兽：是的。我们说，一支信号蚁队一路走着，每到一地，作为它的成员的那些蚂蚁与当地的蚂蚁或者通过直接的接触或者通过交换气味而发生互相作用。接触和气味提供了当地正急需什么的信息。像筑巢、哺育幼蚁等等。当它的供给与当地的需求不能相符时，信号蚁队就会保持内聚力。当它能作出贡献时，它就解散，分出一个新的蚁队前去增援。您现在明白了吗，种姓分布怎样在蚁群中扮演蚁队的总导演的角色？

阿基里斯：我确实明白了。

食蚁兽：您看，这种观察方式怎么会不使人认为信号蚁队的活动是无目的的呢？

阿基里斯：我也认为这样，实际上，我是从两个不同的优势角度观察事物的。从一个蚂蚁眼中看，信号蚁队毫无目的。信号蚁队中典型的蚂蚁只是在蚁群中漫无目标地爬着，没有

什么特定的目的，直到不想往前爬了，就离队了。它的队友们一般不会表示异议，这时，蚁队因瓦解而减员了 但只是减少了一些成员，内聚力并没有削弱。没有计划，没有预测，也不需要进行什么决定正确方向的搜寻活动。但从蚁群的观点看，蚁队正好对用种姓分布语言写出的消息作出了反应。现在从这方面看，它很像是有目的的活动。

蟹：如果种姓分布完全是任意的，那又怎样呢？信号蚁队还会联合和解散吗？

食蚁兽：当然，要是种姓分布毫无意义，蚁群就不能存在很久。

蟹：这正是我的观点。蚁群得以存在是因为它们的种姓分布有意义，这种意义是整体的，从低层次上看不见。如果您不从较高层次方面考虑就缺乏说明力。

食蚁兽：我明白您的意思。但我认为您看事物太狭窄了。

蟹：怎么会呢？

食蚁兽：蚁群的进化已有几十亿年的历史了，少数机制被保留下来，而大多数机制被淘汰了。最后保存下来的那套机制使蚂蚁呈现出我们刚才描述的那种状况。如果您能在一场电影中看到(比人生还要快十多亿倍的)整个进化过程，当然能发现各种机制是作为对外界压力的自然反应而出现的，就像开水中的气泡是对外界热源的自然反应一样。我想您不会不明白开水中气泡的"意义"和"目的"的，是吗？

蟹：是的，但是……

食蚁兽：这是我的观点。不管气泡有多大，都是分子活动的产物。您可以忘掉那"更高层次的法则"。这也适用于说明蚁群和蚁队。从进化这个广阔的角度来看事物，您可以使整个蚁群毫无意义和目的。使意义和目的成为不必要的概念。

阿基里斯：食蚁兽博士，那您为什么对我说您与希拉里大婶谈过话？现在您的意思好像说她根本不会讲话，不会思考。

食蚁兽：阿基里斯，我的话没有自相矛盾。您知道，要在如此宏大的规模上观察事物，我和别人一样有许多困难，所以我发现改变观点要容易得多。当我这样做时，忘记了进化论和从此时此地来观察事物，目的论的词汇就会重现种姓分布的意义和信号蚁队的目的。不仅在我思考蚁群的时候，而且在思考我自己的大脑和别人的大脑时，都是这样。然而如果必要的话，我可以努力记住其他的观点，并使所有这些系统都变得毫无意义。

蟹：进化当然能创造一些奇迹。而您永远不会知道进化下次变出的戏法是什么。例如要是有两支或更多的"信号蚁队"迎面走过，双方都不知道对方也是一支信号蚁队，都把对方当作当地的蚂蚁，这在理论上完全可能。我对此也一点不感到意外。

食蚁兽：不仅是理论上可能。事实上，这类事每天都在发生！

阿基里斯：我想出了一幅非常奇怪的图景。如果那些蚂蚁朝四个不同方向爬动，有些蚂蚁是黑的，有些则是白的，构成了一个十字交叉的有序图案，几乎像……像……

乌龟：也许像一首赋格曲？

阿基里斯：对，就是！一首蚂蚁赋格曲！

蟹：很有趣的想象，阿基里斯，顺便说一句，刚才谈论开水使我想起了茶。谁想再喝些茶？

阿基里斯：蟹先生，我想再来一杯。

蟹：好的。

阿基里斯：您认为，能把这首"蚂蚁赋格曲"的不同的可见的"声部"分清楚吗？我觉得不容易办到……

乌龟：我不要茶，谢谢。

阿基里斯：……找出一个单一的声部……

食蚁兽：蟹先生，我也想要点茶……

阿基里斯：在一首音乐的赋格曲中……

食蚁兽：……如果不太麻烦的话。

阿基里斯：……当所有的……

蟹：不必客气。四杯茶……

乌龟：三杯！

阿基里斯：……声部同时奏出。

蟹：……就来。

食蚁兽：这真是个有趣的想法，阿基里斯，但好像不是每个人都能逼真地画出这样一张图来的。

阿基里斯：这太糟糕了。

乌龟：也许您能回答这个问题，食蚁兽博士。一支信号蚁队从形成到解散是不是总是由同样一组蚂蚁组成的？

食蚁兽：事实上，构成信号蚁队的个别蚂蚁有时会离队，被同种姓的其他蚂蚁所代替，如果该地有这种蚂蚁的话。更经常的情况是，当信号蚁队解散时，没有一个蚂蚁是当初队伍成立时加入的。

蟹：我可以看出，信号蚁队在不断地影响整个蚁群的种姓分布，以此对蚁群的内部需要作出反应，而蚁群的这种内部需要又反映了蚁群所面临的外界环境。因而，食蚁兽博士，正如您所认为的，种姓分布不断地以最终反映外部世界的方式来得到重组。

阿基里斯：但那些中介结构层次的蚂蚁又是如何呢？您刚才说，最好不要根据蚂蚁和信号蚁队，而要根据以其他蚁队为成员的、一层又一层的、直至个别蚂蚁的那些中介蚁队来描

述种姓分布。您说过这是理解可能将种姓分布描述成对外界信息解码的关键。

食蚁兽：是这样。我们正要谈这些。我想给高层次的蚁队起个名字叫"符号蚁队"。注意，这个词的某种意思与平时不同。我说的"符号"是指一个复合系统中的能动的子系统，它们由更低层的能动子系统组成……所以它们完全不同于被动的符号，不是处在系统之外，像字母表的字母或音符那样静静地等待着一个能动的系统来把握它们。

阿基里斯：这未免有点复杂化了，是吗？我根本不知道蚁群还有这样一个抽象结构。

食蚁兽：这是一个很了不起的结构，但是，要存储起能使一个有机体（在任何合理的意义上）成为"有理智的"生物的知识，所有这些层次结构是必不可少的。任何一个掌握语言的系统都有一套相同的基本层次。

阿基里斯：真是糟透了，您的意思是说我的大脑说到底是由一大群四处乱爬的蚂蚁组成的吗？

食蚁兽：嗽，不是这意思。您太咬文嚼字了。最低的层次可能是完全不同的。例如，食蚁兽的大脑就不是由蚂蚁组成的。但当您处在更高一级或二级的层次中，那些成员在其他智力相同的系统例如蚁群中就会有精确的副本。

乌龟：这就是为什么可以把您的脑图，阿基里斯的脑图绘入蚁群，而不是单个蚂蚁脑中的原因。

阿基里斯：谢谢您的恭维。不过，这样一种绘入工作怎么进行呢？比如说，我脑中有什么东西相应于您称之为信号的低层次蚁队？

食蚁兽：我对大脑的活动实在是一知半解，所以无法详细说清脑图的情况。如果我错了，蟹先生请您指正。我猜想，大

脑相应于信号蚁群的东西是神经原的激活，也许是大规模的激活，就像一幅神经原激活的图案。

蟹：我同意您的说法。但您不觉得从我们讨论的目的来看，描绘精确复本工作本身尽管是吸引人的，但却是无关紧要的？依我看好像主要之点在于，尽管我们现在不知道该如何确切地规定这种相应性，但它确实存在。食蚁兽博士，我只对您的观点提一个问题，它关系到在哪一个层次上人们可以相信相应性出现了。您好像认为在大脑中可找到一支信号蚁队的直接复本。而我觉得，只是在您的能动的符号层次中及其以上层次中，好像才肯定会出现这种相应性。

食蚁兽：您的阐述比我的精确。蟹先生，谢谢您阐发了这个如此微妙的观点。

阿基里斯：那么符号能做哪些信号所不能做的事呢？

食蚁兽：这有点像词与字母的关系。词是有意义的实体，由字母构成。而字母本身是无意义的。这就充分表明了符号与信号之间的区别。事实上，这是个有用的类比。只须记住一点，词和字母是被动的，信号和符号是能动的。

阿基里斯：我会记住的。但我不明白，为什么这样强调能动的和被动的实体间的区别。

食蚁兽：原因是您赋予任何被动符号（如写在纸上的一个词）的意义实际上是来源于您大脑中相应的能动符号所含的意义。所以只有当被动符号的意义与能动符号意义相关时才能被正确理解。

阿基里斯：对。您说信号完全是个自主的实体，没有意义，那么是什么东西将意义赋予一个符号——一个肯定是能动的符号？

食蚁兽：这完全与那些符号能引起另一些符号发生反应的方式

有关。当一个符号变得能动时，不会处在孤立状态之中。其实它浮动于一种以种姓分布为特征的介质之中。

蟹：当然，在大脑中没有这样的种姓分布，但相应物是“大脑状态”。在描述大脑状态时，您描述了所有神经原的状态及其一切相互关系和每个神经原激活的限度。

食蚁兽：很好。我们把“种姓分布”和“大脑状态”放在一起讨论，统称它们为“状态”。现在既可以在低层次上也可以在高层次上描述这种状态。有关一个蚁群状态的低层次描述将涉及到对每个蚂蚁的位置、年龄、种姓及其他类似方面作出无比细致的分析。非常详细的描述实际上不可能产生有关这种状态原因的宏观认识。另一方面，有关高层次的描述将涉及对哪些符号能在哪些条件下被另一些符号的哪种结合所触发等问题的分析。

阿基里斯：那么有关信号蚁队层次或蚁队层次的描述是怎么回事呢？

食蚁兽：对这个层次的描述处于低层次与符号层次描述之间。它将包含大量的有关蚁群各处特定位置上实际情况的信息。尽管肯定要比逐个蚂蚁地描述所得的信息要少，因为蚁队是由一群蚂蚁组成，逐队蚂蚁的描述就像是逐个蚂蚁描述的总结。然而，前一种描述还有一些额外的东西是后一种描述所缺乏的，如蚁队之间的关系，各种种姓在各处的供应情况等。这额外的复杂性就是您有权总结的代价。

阿基里斯：我感到比较各层次描述的长处很有意思。最高层的描述好像最富于说服力，它给您展现了一幅最富直观性的蚁群图。尽管奇怪的是，它好像省略了表面看来最重要的特征——蚂蚁。

食蚁兽：但您知道，任凭表面看来怎样，蚂蚁并不是最重要的

特征。众所公认，没有蚂蚁，就没有蚁群；但相应的事物——大脑——却没有蚂蚁也能存在，所以至少从高层次的观点看，蚂蚁是不重要的。

阿基里斯：我肯定，没有哪个蚂蚁会热情拥戴您的理论。

食蚁兽：我从未见过有高层次观点的蚂蚁。

蟹：食蚁兽博士，您给我们描述了一幅多么有趣的反直觉图画。好像是如果您的话不错，那么为了掌握整体结构，在描述它时就必须对它的基本构成忽略不提。

食蚁兽：也许我用个比喻能说得更清楚些。设想您面前有一本查尔斯·狄更斯的小说。

阿基里斯：是《匹克威克外传》，行吗？

食蚁兽：好极了！现在设想您在做下面的游戏：您必须找出一种将字母描绘成观念的方法，以便一个字母一个字母地读这本书时也能懂得全书的意思。

阿基里斯：晦……您是说我每次遇到一个词像"the"时，就得考虑三个确定的概念，一个接一个，没有改变的余地？

食蚁兽：对，它们分别是概念"t"，概念"h"和概念"e"。这些概念每次都像以前那样出现。

阿基里斯：这下就像是把"读"《匹克威克外传》的体验变成了一个难以形容的、令人讨厌的恶梦。不管每个字母使我联想起什么概念，都会成为一种毫无意义的练习。

食蚁兽：就是。从个别字母到真实世界之间没有自然描绘关系。自然的描绘关系出现在更高的层次上，在词与真实世界的组成部分之间。如果您想描述这本书，就不必提字母。

阿基里斯：当然不提！我只描写情节和人物等等。

食蚁兽：这就对啦。您必须对一切基本构成略而不提，尽管有了它们才写出了这本书，它们只是介质，而非信息。

阿基里斯：对。那么蚁群呢？

食蚁兽：在蚁群中，能动的信号代替了被动的字母，能动的符号代替了被动的词，但道理还是一样。

阿基里斯：您是说我不能在真实世界的事物与信号之间建造一种描绘关系？

食蚁兽：您会发现，没法用触发新信号的方式来做到这一点。在任何更低的层次上，比如在蚂蚁的层次上也做不到。只有在符号的层次上，触发性的形式才有意义。例如，设想有一天您正注视着希拉里大婶，正巧我来拜访。无论您看得多么仔细，也可能除了一堆重新组合的蚂蚁之外，什么也没看见。

阿基里斯：我相信正是如此。

食蚁兽：然而，当我从较高层次而不是从较低层次观看时，我看见几个潜藏着的符号显现了，翻译过来就是这样的想法："瞅，那个迷人的食蚁兽博士又来了，多高兴啊！"或者是一些表达这类思想的词。

阿基里斯：这就像刚才我们四个在解读MU图时发现了不同层次那样，或者至少是我们三个发现了不同层次……

乌龟：在我偶然看到的《平均律钢琴曲集》中那幅奇妙的图与我们的谈话倾向中，竟有如此相似之处，真是令人惊奇的巧合。

阿基里斯：你认为这只是巧合吗？

乌龟：当然。

食蚁兽：啊，我希望您现在懂得了希拉里大婶思想中怎样通过操纵由蚂蚁组成低层次蚁队，再由此组成高层次蚁队，再组成信号蚁队，再组成符号蚁队这一系列的活动来产生思想的。

阿基里斯：为什么您把它叫作“符号操纵”，如果符号本身是能动的，那么谁来操纵呢？谁是动因呢？

食蚁兽：这就回到您先前提出的有关目的的问题上。您说得对，符号本身是能动的，然而它们的活动并不绝对自由。所有符号的活动都是由它们所在的整个系统的状态严格限定的。因此，整个系统决定了它的符号相互间怎样触发，所以说整个系统就是“动因”是非常合适的。当符号活动时，系统的状态慢慢发生了转变或更新，但始终还保留着很多特征。就是这种部分永恒、部分变动的系统才是动因。人们可以给整个系统起一个名字。例如，希拉里就是那个可以说是在操纵她的符号的“谁”；阿基里斯，你也一样。

阿基里斯：这是有关我是谁的概念的一个非常奇怪的特点。我不敢肯定自己能完全理解它，但我会给出我的看法。

乌龟：当您在思考自己大脑中的符号时，我们来理解您脑中的符号会是一件很有趣的事。

阿基里斯：我看，这太复杂了。我觉得就是难以想象怎样在符号层次上观察和理解蚁群。我当然能在蚂蚁层次上观察蚁群。而且尽管有些困难，但我还是能想象出在信号层次上观察到的蚁群会是什么样；可是在符号层次上去观察一个蚁群究竟是怎么回事呢？

食蚁兽：只有通过长期实践才能知道这一点。但要是处在我的位置上，就可以像你们在MU图画里看见“MU”那样容易地看到最高层次的蚁群。

阿基里斯：真的吗？那一定是惊人的体验。

食蚁兽：阿基里斯，用一种方法，而且也是您相当熟悉的那种方法就能看到。

阿基里斯：我熟悉的方法？您这是什么意思？我只会从蚂蚁的

层次上观看蚁群。

食蚁兽：也许不是，但蚁群有许多方面与大脑差不多。

阿基里斯：可我也从未见过和研究过大脑。

食蚁兽：那么您自己的大脑呢？难道您不了解自己的思想？它不是您意识的精华吗？除了直接地在符号层次上研究您自己的大脑以外，您还在做什么？

阿基里斯：我从没那样想过。您是说忽略了所有的较低层次，只看到了最高的层次？

食蚁兽：在意识系统方面正是这样。它们只在符号层次上感知到自己，根本没有察知像信号层次这样较低的层次。

阿基里斯：在大脑中是不是也有一些能动的符号不断地在更新自己，以便总是在符号的层次上反映大脑的总状态？

食蚁兽：当然，在任何意识系统中都有代表大脑状态的符号。它们本身又是以符号形式所代表的大脑状态的组成部分。因为意识需要高度的自我意识。

阿基里斯：这真离奇。这就是说，尽管我的脑中每时每刻都在发生狂乱的活动，但我只能用一种方法即在符号的层次上记录下这种活动。我对低层次的活动完全一无所知。这就像从来没有学过字母却能直接通过视觉来阅读狄更斯的小说。我不能想象真会发生这样离奇的事。

蟹：但确切地说，当您只看见"MU"，没看见低层次的"HOLISM"和"REDUCTIONISM"时，这种情况就确实发生了。

阿基里斯：您说得对。我忽略了较低层次，只看到最高层次。我怀疑是否我也忽视脑中所有低层次的含义，只是理解了符号层次的含义。最高层次不包括有关低层次的全部信息，

所以，也不能通过理解高层次而了解低层次的情况，这真是太糟了。但我认为，要是希望最高层次能从最低层次中解出某些信息，这种想法未免太天真了。“MU”图是最显著的例子：最高层次只说出了“MU”，它与最低层次毫无任何联系！

蟹：完全正确(拿起“MU”图，更仔细地看着)。唔，在这幅图中最小的字母有些怪：它们是扭动的……

食蚁兽：让我看看。(仔细审视着“MU”图)我觉得，还有另外一个层次我们都没发现。

乌龟：您说说吧，食蚁兽博士。

阿基里斯：噢，不，那不可能！让我看看。(非常仔细地看着)我知道，你们几个都不会相信的，但这张画的信息就在我们的面前，只是隐藏在深处，它不过是一个词，像咒语那样重复出现，——一个多么重要的词，“MU”！你们知道吗，就与最高层次的那个“MU”一样，我们谁都不会怀疑的。

蟹：阿基里斯，要不是您，我们决不会注意它的。

食蚁兽：我想知道，最高层次和最低层次相同是否纯属巧合？还是某个造物主的有目的活动的结果？

蟹：怎么能判定呢？

乌龟：我觉得没有什么办法，因为我们不知道为什么在蟹先生的《平均律钢琴曲集》中会有这么不寻常的插图。

食蚁兽：尽管我们一直在进行激烈的讨论，但我还是很注意听这首长而复杂的四声部赋格曲。它太美啦。

乌龟：当然是的。嗽，现在就要轮到管风琴曲调了。

阿基里斯：当一段乐曲缓慢下来，在一个音符或和音上休止片刻，然后经过短暂寂静又恢复正常速度时奏出的就是管风

琴曲调，是吗？

乌龟：不是。你说的是"延音"，一种音乐中的分号。在前奏曲中有一个延音，您注意到了吗？

阿基里斯：恐怕我刚才没留神。

乌龟：对了，您还有机会听到延音的，实际上，在这首赋格曲的尾声里还会出现几个延音。

阿基里斯：噢，太好了，您能提前提示一下吗？

乌龟：如果您想听，那当然可以。

阿基里斯：但请告诉我，什么是管风琴曲调？

乌龟：管风琴曲调是通过复调声部中的一个声部（通常是最低的声部）来持续一个单音符，而其他声部则继续按原来各自独立的步骤演奏下去。这种管风琴曲调在G·利斯顿的音符中很细致，您会听到它的。

食蚁兽：有一天我在希拉里大婶那里聊天，发生了一件事，使我想到了您的提议——去观察阿基里斯大脑中的那些能形成有关自身思想的符号。

蟹：快把那件事告诉我们。

食蚁兽：那天，希拉里大婶觉得很孤独，很想找个人说说话。所以，她非常乐意地让我尽情享用那些味道最鲜美的蚂蚁（她对她的蚂蚁总是非常慷慨的）。

阿基里斯：哎呀！

食蚁兽：正当我在仔细观察表现她思想的那些符号时（因为其中有些蚂蚁看上去味道一定很美），那件事发生了。

阿基里斯：是吗！

食蚁兽：于是，我毫不客气地吃了一些属于我正在解读的高层次符号的肥蚂蚁。特别是它们曾属于表达下述思想的一部分符号："你尽管享用那些美味的蚂蚁，请不必客气。"

阿基里斯：哎呀！

食蚁兽：对它们是太不幸了，但对我却很幸运，这些小虫子根本没有想到它们集体构成的一些符号会表示什么意思。

阿基里斯：哎呀！这真是一个有趣的圈套。它们完全不知道自己所干的那些事。它们的活动能被视为一个高层次形式的组成部分，可它们自己当然一无所知。噢，真遗憾，真是极大的嘲讽。实际上，它们忽视了。

蟹：您的话不错，龟先生——那是段多么美妙的管风琴曲。

食蚁兽：我以前从未听见过，但这段乐曲很清晰，一般不会被人忽略。效果极佳。

阿基里斯：什么？管风琴曲已经过去了？如果这样清晰，那我怎么没听见？

乌龟：也许您太全神贯注在说话了，完全没有注意到曲子。噢，真遗憾，真是莫大的讽刺，您实际上会没听见。

蟹：告诉我，希拉里大婶住在一个蚁冢里吗？

食蚁兽：是的，她拥有相当多的财产。这财产以前是属于别人的。这是一段很不愉快的故事。总之，她很富有，与其他许多蚁群相比，她生活得非常奢侈。

阿基里斯：那这怎么能与您先前所叙述的蚁群的共产主义性质相符呢？在我听起来是自相矛盾的。口头上是共产主义，实际上却是骄奢淫逸！

食蚁兽：共产主义是在蚂蚁的层次上实行的。在一个蚁群中，所有的蚂蚁都为谋求共同的食物而劳动，即使有时对它们个体不利，也不会改变这个宗旨，这只是希拉里大婶结构的内在方面。但就我所知，她也许根本没有意识到这种内在的共产主义。绝大多数的人对自己的神经原的活动一无所知，其实他们也许对自己丝毫不了解大脑的活动状况感到

很宽慰，因为人是比较脆弱的动物。希拉里大婶的神经也是比较脆弱的。每当她真正考虑蚂蚁时，她就显得更富于蚂蚁性。所以她总是尽量不去想它们。我真怀疑，她怎么会对自己结构中固有的共产主义社会毫无觉察。她自己是一个坚定的自由意志论者——您知道这是主张自由放任一类的事。所以，她住在豪华的庄园里，至少在我看来不足为奇。

乌龟：我边听边翻阅这本可爱的《平均律钢琴曲集》，刚翻到这页，我发现两个延音中的第一个延音马上就要出现了，最好注意听着，阿基里斯。

阿基里斯：我会的，我会的。

乌龟：这页上也有一张很奇怪的插图。

蟹：是另外一张吗？有些什么说明？

乌龟：您自己看吧(把乐谱递给蟹)。

蟹：啊哈，只是几串字母。让我看看，这是些不同的字母"J""S""B""m""a"和"t"，很奇怪。怎么前三个字母是大的，后三个字母又缩小了呢。

食蚁兽：我可以看看吗？

蟹：怎么不行，当然可以！

食蚁兽：噢，您只注意了细节，而完全忽略了大的方面。实际上，这些字母是"f""e""r""A""C""H"，没有一点重复的。它们开始比较小，后来又变得大些了。阿基里斯，您是怎么看的？

阿基里斯：唔，让我看看，喔，当您把它向右移动时，我看见是一组大写的字母。

乌龟：拼出来是什么？

阿基里斯：噢，是，哈……"J·S·巴赫"(J.S.BACH)。我现在明白了，这是巴赫的姓名！

乌龟：奇怪，您怎么会这样看。我看是一组小写字母，越往右越缩小，拼出来……一个名字……（特别在说出最后几个词时话音放慢了，稍为停顿了一下，突然就像没发生什么不寻常的事那样重新喊出）“fermat”。

阿基里斯：我知道，您老在想着费马。您觉得哪儿都有费马大定理。

食蚁兽：您说对了，乌龟先生，我刚才听到这首赋格曲的一个小小的迷人的延音。

蟹：我也听到了。

阿基里斯：您是说除了我以外人人都听见了？我发现自己太蠢了。

乌龟：好啦，好啦，阿基里斯，别那样想。我肯定，您不会错过“赋格曲”的最后延音的（它很快就出现了）。回到刚才的话题上来吧，食蚁兽博士，您提到关于希拉里大婶资产的先前拥有者有段悲惨故事，您指的是什么？

食蚁兽：先前那个物主是个古怪的家伙，他曾在最有创造性的蚁群中呆过。他的名字叫约翰·塞巴斯蒂安·费马。他是专业数学家，业余音乐家。

阿基里斯：嗷，真是多才多艺！

食蚁兽：正当他的创造力达到鼎盛时期，他却最不适时宜地退位了。那一个炎热的夏日，他外出去晒晒太阳，这时一场百年罕见的暴雨却降临了，倾盆大雨淹没了J·S·费马的领地，因为暴风雨来得太突然，蚂蚁们张慌失措，乱作一团。那儿十年来苦心经营的精密构造毁于一旦。真是一场大悲剧。

阿基里斯：您是说，所有的蚂蚁都被淹死了，这显然会导致可怜的J·S·费马王朝的覆灭？

食蚁兽：其实不是。蚂蚁们最终还是都幸存下来了。它们爬上漂浮在急流中的棍子和圆木。但当洪水退去，蚂蚁们又被冲回原来的家园时，什么都没有了。种姓分布完全打乱了，而且它们自己也没有能力再重建以前那种完好的组织了。它们就像一些一旦损坏便无法修复的东西。我自己也像蚁王的所有人马一样试着重建可怜的费马王朝。我虔诚地拿出糖和奶酪，一次又一次希望，费马会再次出现……（掏出手帕拭眼泪）

阿基里斯：您真了不起！我从没想到食蚁兽的胸襟这样宽广。

食蚁兽：但这些全都无济于事，他已毁了，不能再重建了。然而，后来又发生了一件非常奇怪的事：大约几个月之后，那些曾经是J·S·费马蚁群成员的蚂蚁逐渐集合起来，建立了一个新组织，从此，希拉里大婶诞生了。

蟹：真精采，希拉里大婶与费马是由同样一些蚂蚁组成的吗？

食蚁兽：对，最初是这样，但现在一些年迈的蚂蚁相继死去，其他蚂蚁取而代之了。但仍有许多J·S·费马时代的遗老。

蟹：在希拉里大婶里您还能常找出J·S·费马的古老特征吗？

食蚁兽：不能。它们没有共同点。依我看，也不该有共同点。通常重新整理一组部分来构成“总和”的方法毕竟有好几种，希拉里大婶是对旧部分的新“总和”。请注意，不过是总和，只是特种的总和。

乌龟：说到总和，我想起了数论。在数论中人们有时能把一条公理拆成它的若干组成符号，再以一种新形式重新组合起

来，构成一条新的公理。

食蚁兽：我从未听说过这种现象，我承认对这个领域一窍不通。

阿基里斯：我也没听说过，虽然我对数论颇为精通，如果我不是自我标榜的话。我猜想伯先生正在精心策划一个玩笑，现在我对他已经很了解了。

食蚁兽：说到数论，我不由地又想起了J·S·费马。因为数论是他最拿手的领域。事实上，他已为数论作出了非凡的贡献。而希拉里大婶在任何只要与数学有一点点关系的方面都非常迟钝。她对音乐的爱好也是停留在非常平庸的水平上。可是塞巴斯蒂安的音乐天赋极高。

阿基里斯：我对数论很感兴趣，您能给我们讲讲塞巴斯蒂安的理论贡献的性质吗？

食蚁兽：可以。那么(稍停片刻，呷了口茶，然后继续讲)您听说过福尔米那个臭名昭著的"多次检验的猜想"吗？

阿基里斯：我不太清楚，听起来很熟悉，但我说不清。

食蚁兽：它是个非常简单的概念。福尔米是个专业数学家，但又是个业余律师。有一天他正在研读经典著作《算术》一书，见到书中有这样一个方程：

$$2^a+2^b=2^c$$

他立刻意识到这个方程有无穷多个解，a、b、c。然后他在页边写下了以下这段著名的评注：方程

$$a^n+b^n=c^n$$

只有当$n=2$时，得出a、b、c和n为正整数的解。(因而有无穷多个a、b、c能满足这个方程)不过当$n>2$

时，方程无解。我已发现证明这一条的绝妙方法，可惜页边太小，写不下。

从此以后300年过去了，数学家们一直在徒劳地做着这样两件事：或者是证明福尔米的论点，从而维护他的声誉（虽然他有很高的名誉，但有些怀疑论者认为，他因为从未真正发现他声称已发现的那种证明法而声名狼藉）；或者想找出反例来驳斥福尔米观点：当$n>2$时，整数a、b、c和n能满足这个方程。直到最近，这两方面的尝试都失败了。诚然，这种"猜想"已被证明出n的很多特定值——特别是，所有的n都达到125，000。然而只有约翰·塞巴斯蒂安·费马才证明所有n都是如此。是他找到了为福尔米澄清名声的证明。现在人们都知道这个证明叫"约翰·塞巴斯蒂安的多次证明的猜想"。

阿基里斯：如果它最终给出了准确的证明，与其叫"猜想"还不如叫"公理"更合适？

食蚁兽：严格地说，您是对的，但人们一直这样叫它。

乌龟：塞巴斯蒂安创作了哪种乐曲？

食蚁兽：他很有作曲的天赋。可惜的是，他的最伟大的作品笼罩着神秘的气氛，所以从未发表过，有些人认为，他的作品全装在脑袋里；另一些人更不客气地说他可能从未写出什么东西，只是说说而已。

阿基里斯：那么这些杰作的特点是什么呢？

食蚁兽：它是庞大的先奏曲和赋格曲。赋格曲有24个声部，包括24个不同的主题，每个大调和小调各有一个主题。

阿基里斯：一定很难把24个声部的赋格曲当作整首曲子来听。

蟹：更不用说创作它了。

食蚁兽：我们只是从塞巴斯蒂安对它的描述中得知的。他把这些话写在柏格兹特胡德的管风琴前奏曲和赋格曲的乐谱页边上，写完最后这些话，他就悲剧性地退位了：

我已经创作出了一支绝妙的赋格曲。在这首曲子里，我把24个主音调子和24个主旋律加在一起；构成了一支带有24个声部的赋格曲。但很可惜，页边太小，写不下。

这首没有问世的杰作被称作"费马大赋格曲"。

阿基里斯：噢，这真是难以容忍的悲剧。

乌龟：说到赋格曲，我们正听着的这支赋格曲就要结束了。在接近尾声时，主旋律出现了很奇怪的新变奏。(翻开《平均律钢琴曲集》中的那一页)嗽，我们是听到这儿了吗？一幅新插图，真迷人！(把它递给蟹看)

蟹：看，这是什么？噢，我明白了，它是用大字母写的"HOLISMIONISM"。先是慢慢缩小，后又逐渐放大。这没什么意义，因为它不是个词(把它递给食蚁兽)。

食蚁兽：看，这是什么？噢，我明白了，它是用小字母写的"REDUCTHOLISM"。先是慢慢缩小，后又逐渐放大。但这没有什么意义，因为它不是个词(把它递给阿基里斯)。

阿基里斯：我说出来你们都不会相信的，这幅图其实是由两个"HOLISM"组成，并且字母越往右越缩小(把乐谱又递还给乌龟)。

乌龟：我说出来你们都不会相信的，事实上，这幅图是由"REDUCTIONISM"一个词组成的，并且字母越往右边越大。

阿基里斯：这回我终于听到了主旋律的新变奏了！龟先生，谢

谢您的指示，我想我终于能掌握听赋格曲的方法了。

D·R·霍夫施塔特

反　　思

灵魂大于它部分的集合吗？前面对话中参加者们好像对此问题有分歧。然而，他们肯定的和一致赞同的是，由个体构成的系统的集体行为可以有许多惊人的特征。

许多人读了这篇对话后便联想起了国家的那种表面看来有目的的、自私的、以生存为主导的行为。这种行为以某种方式源自其公民的习俗和体制：他们的教育体制、法律机构、宗教、能源、消费形式和期望的层次等等。当一个牢固的组织是由不同的个体组成——尤其当对该组织的贡献不能追溯到较低层次的具体个人身上时，我们便倾向于把这个组织当作一个高层次的个体，并且经常用拟人的词汇来谈论它。一张报纸登载了一篇有关恐怖集团的文章，文章把这个恐怖集团说成是“手法极其高明”。人们常说苏联“渴望”全世界承认它的威力，是因为它对西欧“患有”“长期的自卑综合症”。这些众所公认的比喻表明，人们多么喜欢对组织作拟人化的描述。

作为组织成员的个体——秘书、工人、公共汽车司机、行政官员等等，有着自己的生活目标。这些目标有可能与他们所构成的高层次实体的目标发生冲突。但是这种组织可以采取一种（政治学专家将视之为阴险狡猾的）手法，利用个体的自大和自尊等要求来吸取和利用他们的生活目标，以达到为自己服务的目的。从许多低层次的目标中产生了一般高层次的总冲量。

它将所有低层次目标都包摄和裹挟在自身，以求得永存。

因此，乌龟反对阿基里斯把自己比作蚂蚁，而赞成阿基里斯在适当的层次上将"自己绘成"一张蚁群图的做法也许不算很蠢。同样，我们有时自问："是中国会怎样？与是美国有什么不同？"这样的问题有意义吗？现在我们还是考虑一下"是"一个国家有没有意义。一个国家有思想或信仰吗？这都归结于一个问题，一个国家是否有希拉里大婶那样的符号层次。我们也许不说一个系统"有一个符号层次"，而说"它是一个表象系统"。

在本书中这个"表象系统"的概念极为重要。需要给出较为确切的定义。所谓"表象系统"是指组织起来的、能动的、自我更新地"反映"发展着世界的结构体。一张画不管怎样富于表现性，都不能算作表象系统，因为它是静止的。奇怪的是，我们也把镜子本身排除在外，尽管可以证明镜子中的一组影像是与世界完全相符的！镜子不能成为表象系统的原因是双重的。第一，镜子本身并不区分不同物体的影像——它反映的是宇宙，但根本看不见范畴。事实上，一个镜子只映出一个影像——在观看者的眼中，镜子的唯一影像分解成许多不同物体的"分离"影像，镜子不会感知，只会反映。第二，镜子中的影像不是一个有自己"生命"的自主的结构。它直接依赖外部世界。假如灯被关掉，影像也就消失了。一个表象系统即使在与它所"反映"的现实中断关系之后也能工作(虽然现在说的这种"反映"还不是一个很充分说明问题的比喻)。独立的表象结构现在如果不是以反映世界真正发展的方式，至少也是以反映世界可能发展的方式继续进化。实际上，一个优秀的表象系统能为各种可被合理预见的可能性提供发展的基础。它的内在模式在隐喻的意义上被定义为对"再现心灵"的反思，这些模式进入叠加状态，每个都带有对可能发生之事的联想的主观判断。

简言之，表象系统是建立在范畴之上的；当有必要提高或扩大其内在范畴网络时，它就把输入的材料筛选出来纳入这些范畴之中；它的表象或"符号"依据自己的内在逻辑相互作用。这种内在逻辑的作用虽然不参照外界的作用，但却构成了一个十分可信的外界作用模式，因而能使那些符号与它们所反映的世界非常"相符"。所以电视不是一个表象系统，因为它不加区别地把小光点投向荧光屏而不考虑它们代表什么东西。荧光屏上的图案没有自主性——不过是"远处"事物的被动复制品。相比之下，能"观看"景象并告诉你那景象是什么的计算机程序更近似一个表象系统。有关计算机想象力最先进的人工智能工作还没有突破。程序可以观看一幅景象，它不仅能告诉你这幅景象中有什么，而且还能告诉你它的原因可能是什么，结果可能是什么。这就是我们所指的表象系统。在这种意义上说，一个国家能是一个表象系统吗？一个国家有符号层次吗？这些问题将由读者自己去思考。

"蚂蚁赋格曲"的最重要概念之一是"种姓分布"或"状态"。有人认为它是决定有机体未来的动因。这好像与下述思想相悖：一个系统的所有活动都取决于基本的法则，在蚁群中取决于蚂蚁的法则，在大脑中取决于神经原的法则——无论在哪方面最终都取决于粒子运动的法则。有"下降的因果关系"吗？严格说来，思想可以影响电子的运动方向吗？

在威廉·卡尔文和乔治·奥杰曼著的《大脑内部》一书中，对一个神经原的激活提出了一系列的问题："是什么引起激活的？"是什么引起钠通道打开的？(钠通道的作用是让钠离子进入神经原，当钠离子高度集中时，便触发了神经介质的释放，神经介质从一个神经原流向另一个神经原，构成了神经原激活的基本条件。）答案是：钠通道对电压很敏感，当它们受到一定

强度的电压脉冲冲击时，就立即被打开了。

“但是电压最初是怎么形成的，怎么能达到一定强度……引起一系列叫作脉冲的东西？”回答是：只是由神经轴突的各个“交点”把这种高电压逐段传递过来的。于是这个问题又发生了转化。这次是问：“第一个交点上的第一个脉冲是哪里来的？那个电压从哪里转变过来的？在这个脉冲形成之前又有什么？”

对大脑内部的大部分神经原来说——“内神经原”是指不是由传感输入信号馈给，而是由其他神经原馈给的神经原——答案是：第一个交点的电压变化是由来自其他神经原的神经介质的整个效应引起的（我们可将那些神经原称作“溯流”神经原，其实这样称是不正确的，它意味着大脑中神经活动流向是线性的，只有一个方向，像河里的流水那样。事实上，神经流向形式通常不是线性的，而是不断地构成循环，与河流截然不同）。

这样，我们似乎陷入了恶性循环之中，就像陷入了鸡和鸡蛋孰先孰后的争论之中。要是问：“是什么引起神经原激活的？”回答是：“其他神经原的激活！但这仍没有回答真正的问题。”为什么是这些神经原而不是其他的神经原引起的？为什么是这种恶性循环而不是大脑中另一部分的另一个神经环路引起的？为了回答这个问题，我们必须变换层次来谈论大脑与它所译解的观念之间的关系。那么，这又要求我们讨论大脑是怎样译解或表象它对外界的概念的。既然我们不想在本书中对这类问题作详细的理论探讨，那就只谈一个有关的、比较简单的概念。

假定有一个错综复杂的、时而分叉时而聚合的多米诺骨牌网络，假定每块骨牌底下都有一种时间差很小的反弹力，能使它倒下5秒钟之后又站起来。通过将这张网排成各种图案，其实可以将这个多米诺系统编成程序来进行数字运算，就像一台计

算机那样。各条通路会完成计算的各个部分，并且构成了分支的环路（请注意，这幅图景与大脑中神经原网络相去并不太远）。

人们可以想象一个把整数641分成其质数因子积的“程序”。“为什么这块多米诺骨牌没有倒下？”你指着那块你注视了很久的骨牌提出这样的问题。在一个层次上的答案会是“因为它前面的那块没有倒下”。但是这种低层次“解释”的依据也是需要给予解释的。我们真正想得到的——事实上是唯一令人满意的答案——是在程序概念层次上的答案：它永远不会倒下是因为它所处的那块多米诺骨牌只有当找到一个公约数时才会倒下，而641是质数，不会有公约数的。所以说，这块多米诺骨牌从不倒下的原因与物理学或与多米诺链毫无关系，只是因为641是质数。

可我们因而就会承认，高层次规则对整个系统的形成和控制作用已超过低层次规则了？不，这只说明，一种有意义的解释需要使用高层次的概念。这些多米诺骨牌当然不知道自己是一个程序的组成部分，它们也没必要知道，就像钢琴键不知道，也没必要知道你弹奏的是哪支曲子一样。如果它们真知道的话，会是多么不可思议！你的神经原也根本不知道你用它们进行的现在的这些思想活动，而蚂蚁对它们是蚁群宏大规划组成部分这一点当然也一无所知了。

在你脑中会提出一个更深入的问题：“程序和多米诺链——其实是多米诺骨牌的构造——的存在取决于哪个层次上的哪些法则？”为了回答这个及其由它所不可避免地引起的许多问题，必须越来越向后回溯，回溯到社会存在的全部原因，乃至回溯到生命的起源等原因。最好是把这些问题都撇在一边，只留下一个理由：641的质数性。我们更喜欢这种简洁的高层次说明，把长期的观点作为过去的东西而不予考虑，只注意现在或永恒。但如果我们想追溯事件的根本原因，就不得不使用

道金斯或乌龟所描述的那种还原论观点。的确，最终，我们会回到物理学家那儿，他会告诉我们万物的最初原因是宇宙起源的“大爆炸”理论。然而，这还不能令人满意，因为我们想要的是用人所熟悉的概念来解释的答案。幸好自然是多层次的，有可能提供这种答案。

我们提出是否一种思想能影响电子的快速运动过程。读者可以很容易地想象一个我们脑中本来没有的形象，有一个全神贯注的“巫师”用皱起的眉头发射出“阴府式的能量波”(或别的东西随他怎么叫)，这些波射向一个物体——比如一个滚动着的骰子——决定了它的停顿方式。我们不相信这类事。我们不相信还会有某种尚未发现的“心理磁力”，概念通过它能够“抵达”粒子的轨道，并能通过某种“语义的潜能”来使粒子脱离当代物理学预见其所必经的轨道。我们谈论的是另一回事。这个问题关系到力从何而来，也许关系到恰当用词，关系到“原因”一类词的日常用法与科学用法怎样协调的问题。那么，参照较高层次的概念像“信念”、“意欲”等来解释粒子的轨道合理吗?读者也许会觉得，这种说话方式很实用。就像进化论生物学家常喜欢用“目的论速写法”把他们的概念缩减到直觉看来比较合理的规模那样，我们觉得研究思维机制的人必须善于转换使用“还原论”语言和“整体论”语言，在后一种语言中，整体确实对其部分产生明显的影响，并确实具有“下溯的因果关系”。

在物理学中，当一个观点改变时，有时就会采取另一些法则。想想公共游乐场的旋转木马。当人们在大圆筒式的内墙里靠边坐定后。圆筒开始旋转，边旋转，它的底部边豁开，好像一个巨大的开罐头刀从底部打开了罐头。人们悬在半空，背部靠所谓的离心力贴着内墙。假如你骑在马上，想把一个网球扔给坐在对面墙边的朋友，就会看见球歪歪斜斜飞出跑道，也许

会像那种扔出去会返回的飞镖一样回到你面前。当然，这很简单，因为在球（以直线的形式）穿过圆筒时，你也在绕转。但如果你没意识到你在旋转，你也许会给使你的球偏离预定目的地的奇怪偏转力编造一个名称。你会认为它是地球引力的某种奇异形式。这一点将被下述观察结果所确证：当地球引力发生作用时，这种力量对两个质量相同的物体的作用相等。有趣的是，这简单的观察结论——“虚构的力”和引力容易混淆——成为爱因斯坦的伟大的相对论的核心。这个例子意味着参照系的改变会导致知觉和概念的改变——感知因果关系方式的改变。假如爱因斯坦承认这一点，那我们也应该承认这一点！

我们不想再对整体层次和部分层次间来回摆动的观点转换议论不休，只想介绍一些能吸引读者对这些问题作进一步思索的术语。我们已经比较了“还原论”和“整体论”。现在你们可以明白了，“还原论”是“上溯因果关系”的同义词。“整体论”是“下溯因果关系”的同义词。这些概念与空间中不同规模的事件如何相互限定有关。它们在时间维度中有相对应的概念：与还原论相应的是不考虑有机体“目标”的从过去预示未来的观点；与整体论相应的观点是认为，只有无生命物才能这样被预见，而对有生命物来说，目的、目标和愿望等是解释它们活动的根本因素。这个观点，通常称做“以目的为主”，或“目的论的”，也可以称作“目标论”。它的对立理论可称作“预示论”。于是，还原论在时间方面的对应理论就是预示论。而整体论在时间方面的对应理论则是目标论。预示论是这样一种学说：在决定现在流向未来的方法时，只须考虑“向上流动的”事件，而无须考虑“向下流动的”事件。与之相反，目标论视有生命物体朝未来目标行进着的，因而，在某种意义上它认为未来事件在时间上是向后投射因果力的，或是追溯性的。我们可以称之为“追溯性因果关

系"，它是整体论的"内溯因果关系"在时间方面的对应概念。在整体论中，原因被看作是"向内"流动的（从整体流向局部）。把目标论和整体论结合在一起，你就会猜测到——灵魂论！把预示论和还原论结合在一起，你就得出——机械论。

为了总结起见，我们可以划一个小表：

硬科学家	软科学家
还原论	整体论
（上溯因果关系）	（下溯因果关系）
＋	＋
预示论	目标论
（上行因果关系）	（下行因果关系）
＝机械论	＝灵魂论

既然我们已经想象出了这么些巧妙的用词，那就继续讨论下去吧。另外一个有关大脑活动的比喻给我们提供了一个新观点。这个比喻把大脑比作"会思考的风吹编击乐器"。设想一套编击乐器像一个灵活的装置，上面吊着许多玻璃"铃铛"，这些铃铛像大树枝上长着小树枝，小树枝上长着树叶那样悬挂着，当风吹来时，许多铃铛就颤动起来，于是整个结构在各层次上都逐渐地改变了。显然，使那些小玻璃铃铛颤动的原因不光是风，还有这套乐器的自身状态，即使只有一个玻璃铃铛悬挂着，它的吊绳的盘旋缠绕也会使这套乐器颤动起来，其强度不下于风吹。

就像人们做事"出于自己的意志"那样，这套编击乐器似乎也有自己的意志。什么是意志？这是一个复杂的经过长期进化而来的内部构造，它能将拥护未来某种内在构造反对另些构造

的倾向译解出来。眼下这种最低级的风吹编击乐器就是如此。

但这合理吗？一个风吹编击乐器会有欲望吗？这种乐器能思考吗？我们不免想象一下，给它增加一些特性。假定在这套编击乐器边上安装一台鼓风机，它的位置是受该乐器中某一枝条角度控制的，它的叶片旋转速度则由另一个枝条角度控制的。现在这套乐器就能对周围环境有所控制了，就像有一只被许多很小的似乎无关紧要的神经原控制的大手一样：这套乐器在决定它自己的未来时可以发挥更大的作用。

再假定许多枝条都控制着鼓风机，每个枝条控制一台鼓风机。当(自然的或鼓风机打出的)风吹来时，一群玻璃铃铛就会微微颤动，然后再传送到乐器的其他部分，扩散开来，逐渐使枝条摆动，改变了原来的角度，使乐器形成了影响鼓风机位置和风力强度的新状态，这些变化反过来又引起乐器更多的反响。现在，外部风的作用和乐器的内在状态的作用非常复杂地交错在一起，实际上，从概念上很难分清它们。

设想在一个房间里有两组编击乐器，每组都通过向对方吹送微风来影响对方。谁能说，把这个系统分解成两个自然部分是有意义的呢？也许，观察这个系统的最好方法是从高层次枝条着眼。在这种情况下，在两组乐器的每一组中都有 5 个或10个自然的部分——也许下一个层次的枝条是最好的观察单位。在这种情况下，我们可以看见每组乐器有20个或更多个自然的部分。选择哪种观察角度不过是权宜之计。在某种意义上说所有部分都相互作用着，但也可能发现有两个部分，在空间中或在连贯的活动中保持独立，例如那种微微的颤动也许在某个区域停顿下来了，于是我们就可谈论一种与众不同的“有机体”。但请注意，这整体事物是如何仍然可以用物理学来解释。

现在我们可以假定一支机械手的运动由二打高层次枝条的

角度操纵。这些枝条当然与整个乐器状态密切相关。我们可以设想乐器状态以一种奇异的方式控制这只机械手的运动——告诉这只手该拣哪个棋子，并在棋盘上移动。如果它总是拣适当的棋子，走合乎逻辑的棋步，那不就是绝妙的巧合吗？如果它每走一步都是高着，那不更是奇妙的巧合吗？几乎不能这样认为。假如这种情况出现了，那正是因为它不是巧合，而是因为乐器的内部状态有表象力。

这里我们不必再去详细描述这些奇特的颤动结构怎么会存储那些观念的，这情形使人想起一种叫颤杨的树木。读者必须把握的是一个对外界刺激和自身内在结构各层次特征作出反应的系统所具备的那些潜在的微妙性、复杂性和自我包容性。

要将这样一个系统对外界的反应与其自我包容的反应区分开来几乎是不可能的。因为最小的外部干扰也会引起无数微小的相互关联的事件，会导致连锁反应。如果你把这看作是系统对输入的“感知”，那么很清楚，它也是这样来“感知”自己的状态的。这里自我感知与对外界的感知没法区分。

不能预先得出结论说，有观察这样一个系统的高层次方法存在。也就是说，没法保证我们能把这种编击乐器状态译解成一组连贯的英语句子来表达对这个系统的信念，例如，包括对一套下棋规则（以及怎样下好一盘棋！）的信念。然而，当这样的系统通过自然选择而进化，就会形成某种原因，使有些系统幸存，使其他大多数系统被淘汰：有意义的内部组织使该系统去利用环境和控制环境，至少能部分做到这样。

在风吹编击乐器、假设的有意识蚁群和大脑中，有着不同层次的组织。在风吹编击乐器中，这种层次即是指那些悬垂枝条的不同位置层次，最高层的枝条是对乐器总体状态最简洁、最抽象的总结，而那上千个（或上万个）颤动着的单个铃铛则是对乐

器状态的完全凌乱的、间接的、然而却是具体的和局部的描述。在蚁群体中，有蚂蚁、蚁队以及各种层次的信号蚁队，最后是种姓分布或"蚁群体状态"——这也是对蚁群的最深刻然而却是最抽象的看法。像阿基里斯所不解的那样，它非常抽象甚至根本不提及蚂蚁本身！在大脑中，我们只是不知道怎样去找到这种高层次结构，能使人用英语读出存储在大脑中的那些信念。更确切地说，我们只是要求大脑的拥有者告诉我们，他或她相信什么！但我们找不出从生理上确定在哪儿或怎样对信念进行编码的方法。

在上述三个系统中，存在着各种半自动的子系统，每个都代表着一个概念，各种各样的输入刺激能激活某些概念或者符号。请注意，这个观点认为不存在那种能观察所有活动和"感觉"该系统的"内眼"；相反，该系统状况本身表现了这种感觉。否则，那种据说能用这种"内眼"观察内心活动的传奇式的"小人"就得有更小的"内眼"，而这又导致需要更小的"小人"和更小的"内眼"——总之，这会引起最糟、最蠢不过的恶性循环。相反，在这种系统里，自我意识来自于复杂地交织在一起的对外部和对内部刺激所作出的反应。这种形式说明了一个普遍的论点："心灵是被大脑所感知的一种形式。"这或许未免落入循环，但这种循环既不是恶性的，也不是自相矛盾的。

与那种感知大脑活动的"小人"或"内眼"功能最接近的是自我符号——模拟整个系统的一个复杂的子系统。但自我符号不能通过自己所有的小符号（包括它自己的自我符号——这显然会导致恶性循环）来进行感知，而是与普通的（非反思的）符号联合起来构成系统的感知功能。感知存在于总的系统的层次中，而不是存在于自我符号的层次中。如果你想说，自我符号感知到了某物，这不过是相当于在微观胞间的化学生物遗传密码单

位层次上说，一只雄蛾感知到了一只雌蛾，你的大脑感知到了你的心率。

最后要指出的是，大脑需要这种多层次的结构。因为它的机制必须非常灵活以便适应无法预测的，变化多端的世界。刻板的程序会迅速被摈弃。专用于捕捉恐龙的办法没法适用于捕捉长满长毛的猛犸象，对饲养家畜或造地铁就更没有用处了。一个有理智的系统必须能再造自己——能以较为深刻的方式静坐下来，权衡利弊，然后重新组织起来。这种灵活性只要求最抽象的机制保持不变。一个多层次的系统在其最表面的层次上可以有专门适合特殊需要的程序(比如：下棋程序，捕捉长毛猛犸象的程序等等)，而在逐渐深入的层次上，则相应有更抽象的程序。这样，既能控制外部世界，也能控制内部世界。属于这类较深刻程序的有，识别的模型，评价对立证据的模型，决定在数个引人注目的子系统中哪个应得到优先对待的模型，确定怎样把当下所感知的情境看作未来也许类似的场合中可能重现的模型；决定两个概念是否类似的模型；等等。

有关这类系统的进一步描述将使我们深入到认识科学的哲学和技术领域之中，我们不打算作这种深入的探讨。

12 大脑的故事

阿诺德·朱博夫

I

从前，有一个好心肠的年轻人，他有许多朋友，又很富裕。他得知自己除了神经系统外全身都腐烂了。然而，他渴望生活，他热爱生活的感觉。因此，当很有本领的科学家朋友向他提出下面的建议时，他表示了极大的兴趣。

“我们把你的大脑同你的腐烂了的身体分离开来，将大脑放在一种经过特殊处理的营养液中。然后，再将你的大脑接通一台机器，这台机器能够在你脑中诱发任何形式的神经活动，它能给你带来各种完整的感觉，这样，你的神经系统的活动就能够产生或存在。”

为什么在最后一句中要把产生和存在分开来说呢？这是因为尽管所有这些科学家认为确实有这样一种理论，即“感觉的神

经理论"，但他们对这一理论的具体表述不尽相同。他们都知道，无数事实证明，大脑的构造即其活动形式会使人产生某种特殊的感觉。因此，他们认为，最终控制人的特殊感觉(无论它是否存在和像什么)的决定性因素是神经系统的状态，说得再具体一点，也就是(经周密研究而发现的)与意识的各个方面密切相关的大脑的某些区域的状态。那些科学家正是由于相信了这一点，才向他们的年轻朋友提出这样一个建议的。尽管他们中有人认为感觉就是神经活动本身，也有人认为感觉是由神经活动引起的，但这都无关大局，因为他们现在的当务之急是只要小伙子的大脑还活着，在他们的控制下发挥其功能，因而能让他享受生活中的种种感觉，就像他活着时信步漫游，一路上遇到各种有趣的经历那样。原先自然地产生的神经活动，现在可以由他们用人为的方式激发起来。要是他真往池塘中的一个冰窟窿里张望，那么，里面的客观事物会使他体会到梭罗① 所描写的那种情景："鱼儿幽静的客厅里充满了仿佛是穿过一扇毛玻璃的窗户照进来的柔和的光，明亮的铺沙地板同夏天时一样。"小伙子的大脑被泡在装有营养液的瓶子里，孤零零的，没有身体听它使唤，根本到不了梭罗所描写的那个池塘。但要是它真的来到池塘中，他就能够体会到水下那种美丽的景致。

于是，小伙子同意了朋友们的意见，期待着这一方案的实施。一个月以后，他的大脑被保存在装有温暖的营养液的瓶子里。小伙子的那些科学家朋友们，不停地在雇佣的病人身上研究哪一种神经活动形式是对愉快情景的自然神经反应。他们还通过一台复杂的电极机，在这位朋友的大脑里不停地单单诱发这些神经活动。

---

① 梭罗(1817—1862)：美国超验主义作家。——译者

接着，糟糕的事情发生了。一天晚上，值班人喝了酒，醉醺醺地闯进了放着大脑的那间房子，走近放着大脑的瓶子，身子一歪，手就伸进了瓶子，这可怜的脑子一下子分成了两半。

次日清晨，那些科学家朋友得知后非常难过。他们最近几天发现了一些新的感觉，正准备把它们植入小伙子的脑子里。

“如果我们把左右半球重新修复，要有整整两个月的时间才能恢复到允许植入新的感觉。当然，他是不会知道这些的，但我们清楚得很！不幸的是，我们都知道，分成两半的脑子不可能具有合成整体的脑子具有的那种神经结构。因为在整体的脑子中感觉从这一半球传向另一半球的刺激，在一旦分成两半的脑子里，是不可能传递的。”

听了这番话之后，有人灵机一动，想出了一个点子。为什么不试一试下面的方法？我们可以用一些非常细的电化线，将其两头同神经原的突触连接起来，这样，这些线就能接受和释放这些突触的神经刺激，把在左右两半球的各个神经原重新连接起来。提出这个建议的人叫伯特，他说：“这一来，那些本来从一半球向另一半球传送的神经刺激通过电化线仍然可以继续传送了。”

大伙听罢，齐声喝彩，因为制造这样的电化线系统很容易，一个星期的时间就足够了。可是，一位名叫卡桑德的人却不那么乐观。他说：“我们都知道，我们要植入大脑的那些感觉都曾试图给过那小伙子。也就是说，我们都这样或那样地接受了感觉的神经理论。根据这个我们都接受的理论，只要你保持大脑的活动形式，你就可以任意改变具有功能的大脑所处的环境。我们这样说是什么意思呢？让我们看看吧。一个感觉的产生有各种各样的条件，举例说，这个感觉就是三个星期前让我们的朋友所感觉的看池塘冰窟窿的情景。通常这些条件是，大脑必

须长在一个真正的身体内，这个身体必须亲临那个真正的池塘，这个池塘会使人产生我们真已给过他的那种神经活动。而我们只给了他那种神经活动，并没有提供这种活动必备的另外那些环境条件，因为他现在没有身体，还因为我们都相信，问题的关键不在于这些环境，而在于神经活动本身。我们相信，所谓环境条件对于获得某种感觉来说，实际上无足轻重，即使它们对正常情况下产生感觉来说确实是重要的。如果哪一个人可以像我们这样办，可以规避产生冰窟窿上那种感觉的外在条件的正常必要性，那么，这些条件不再是必不可少的事。这说明，在我们关于感觉的看法里，从原则上讲，这些条件对于获得感觉这个单独的事实从来不起什么作用。

"你们建议用极细的电化线将左右两半球连起来，其实是把我们的朋友获得感觉所必需的另一个正常条件看作是不重要的。也就是说，你们的意思同我谈到神经活动环境时候的意思基本一致。不过，你们谈的是关于大脑左右两个半球的近似条件。你们的意思是说，两半球的相互连接，在一般情况下，对于感觉的产生是必要的条件，但是，如果这种近似的联系被中断以后，在特殊的情况下，即在你们将要用电化线连接它们的情况下，感觉还是能够照常产生的！你们无非认为，近似条件对感觉产生并不是必不可少的。但是，即使是把被分裂的大脑中的整个大脑神经系统丝毫不差地重新修复起来，整个大脑的感觉是不是就能产生呢？会不会有可能产生相反的结果呢？在形成某特定完整大脑的感觉过程中，近似条件不光是不能避免的，而且也许是获得完整大脑感觉的绝对的条件和原则，难道不是这样吗？"

但是，卡桑德的担心并没有引起人们的关注。对他的问题，典型的回答是："这两个该死的两半球会不会知道它们之间已不

是通常那样连在一起，而是用线路连在一起的，换句话说，主控语言、思想或其他功能的大脑结构会不会知道这一情况？自己的脑子在外部观察者看来是个什么样子，这一情况我们的朋友在感觉快乐时怎么会关心呢？——还不是和泡在温暖的溶液中一个样？不管大脑两半球是连着的还是分开的，只要他能够感觉快乐，那就行了。要是我们把大脑两半球和嘴巴连在一起，我们的朋友就会说他觉得很快乐。"大伙越说越激动，卡桑德只是在一旁小声嘀咕，说感觉的某个方面"或类似的方面"可能会中断。

可是，就在大家为接通线路而忙了一阵之后，有人提出了反对意见，这项工作真的停下了。他指出，当大脑合在一起正常工作时，神经的刺激从一半到另一半，简直不需要任何时间。但是，要是大脑两半球用线路来连接，那么，刺激传递的时间比原来要长，而在这两半球的其他部位，刺激传递的速度仍然照常，这样，整个大脑的活动就会受影响，仿佛在某个部位发生了怠工。当然，我们不可能确切地说出正常传递速度是多少，但你总会觉得整个大脑的活动受到了影响。

大伙都觉得这个意见提得有道理。这时有一个几乎不懂物理学的人建议用无线电信号来代替线路。我们可以在两半球的表面各装上一个刺激"脉冲管"，它能将脉冲传送给另一半球的裸露而与其分离的神经原，而且也能从这些神经原接受另一半球打算传送给它的任何脉冲。然后，再把脉冲管装进特殊的无线电收发器里。当脉冲管收到从一个脑半球的神经原发给另一个脑半球的神经原的脉冲时，就把这个脉冲用无线电传递给装在另一个脑半球上的脉冲管。此人还异想天开地说，我们完全可以把大脑的两个半球分别装在两个瓶子里，尽管各在一方，它们还是能像在一起时那样感觉如一。

此人认为，用无线电来代替线路有一个优点，因为，无线电波的传递需要时间很短，而通过线路，脉冲的传递所花时间更长些。但是，很快就有人纠正了他的这种说法。因为，无线电系统会有时间间隙的障碍。

尽管如此，脉冲管的想法启发了那个叫伯特的人。他说："大伙想一想，如果我们在脉冲管里输入它原来要通过无线电接收的脉冲，那么，我们就不再需要无线电收发器了。这样的话，就不再需要把脉冲管与无线电收发器连在一起了，而是同'脉冲程序机'连在一起了。这个装置会将你以前给它的任何脉冲程序反复播放。这一方法的妙处在于，进入一脑半球的脉冲不再需要部分地由来自另一脑半球的脉冲真正地激活起来。这样，也就不必等待传送了。装有程序的脉冲管能与大脑其他部位的神经活动配合协调，时间的控制完全与两半合在一起时一样准确。这样就可以很容易将两半球分装在两个瓶里——也许一半就放在这间实验室里，而另一半则放在城市另一头的实验室里，我们可以用各个实验室里的设备来对半个大脑进行试验。这样一来，一切都变得容易多了。还有，实验室里的人员也可以增加。许多人都想参加我们的研究工作，由于名额有限，他们一直在纠缠着我们，执意要加入我们的研究队伍。"

然而此时此刻，卡桑德变得更加担心了。"我们已经排除了近似条件。现在，我们又要排除通常感觉的另一个条件，即实际的因果联系。假定你对感觉产生的必要条件了如指掌。这样，有了所谓的脉冲程序以后，要形成大脑整体的感觉，在一个脑半球中的那些脉冲无需再成为在另一个脑半球中实现大脑整体的感觉的真正原因。可是，其结果是依然存在着大脑整体感觉的事实呢？还是你在排除这一条件的同时也排除了真正获得大脑整体感觉所需的绝对原则与根本条件？"

对这个问题的回答又与先前一样。神经活动怎么会知道哪些脉冲是通过无线电波传送的，哪些是通过脉冲程序机输入的呢？这个事实对主控语言、思维和其他意识功能的神经结构来说，完全是外在的，怎么能用它们来证明这一事实呢？它肯定不能机械地显示出来。除了时间间隙困难被克服之外，用程序或是用线路，其结果难道不是一模一样吗？当嘴和大脑被连在一起时，不管这张嘴所报告的是由程序产生出来的感觉，还是通过无线电波的传递产生出来的感觉，不都是一样的快乐感觉吗？

既然两个脑半球现在单独进行神经活动，那么把两个脑半球的没有因果联系的脉冲加以同步协调是否有意义？当提出这个问题时，立刻有人想出了新点子。既然每个脑半球事实上会从另一个脑半球接受产生某种感觉时本来要接受的所有脉冲，而且，接受这些脉冲同其他的脉冲在时间上也会配合协调，此外，两个脑半球完全可以独立达到这一理想效果，那么，被卡桑德认为是至关重要的“同步条件”是不是还有必要非满足不可呢？人们会说：“脑半球怎么能有所知，当一个脑半球开始工作时，另一个脑半球怎么能根据外在观察者的时间来指出它是在何时开始工作的，对每个脑半球，我们除了说另一个脑半球的工作状况正常之外，还能说些什么？如果他们在某一日在某一个实验室里做了一半的感觉实验，而在另一日又在另一个实验室里用另一个脑半球继续做下一半感觉实验，这又何妨？整个感觉实验活动还是正常进行。感觉照常产生。如果把嘴和大脑连在一起，那么，我们可怜的朋友就会把他的感觉告诉我们。”

关于是不是要把卡桑德称为“拓扑学”——即是否将两个脑半球放在面对面的空间关系上的问题——大伙讨论了一番，但又忽视了卡桑德的警告。

II

十个世纪已经过去了。但人们对这项伟大的工程依然很着迷。当然，银河系到处都是人，他们的技术也得到了日新月异的发展。数以亿万计的人们都想参加这个“感觉大输入”的活动，这既是追求刺激也是出于责任感。当然，在这个愿望的背后，人们还是相信，通过对脉冲进行程序处理，仍然能够让人获得各种各样的感觉。

但是，为了让想参加的人都能参加这个伟大的工程，曾被卡桑德称为产生感觉的必要“条件”在表面上已经有了很大的变化（实际上，从某种意义上讲，这些条件比先前更加保守了。因为，正如我待一会将要解释的，像“同步”这样的条件又恢复了）。前面我们已经讲过，左右脑半球被分装在两个瓶里。而现在，每一个单独的神经原都被分别装在一个瓶里。因为一个大脑有几十亿个神经原，这样，我们每一个人有幸保管一个装有一个神经原的瓶子。

为了对问题有一个清醒的认识，我们必须再回到十个世纪以前，再看看由于越来越多的人也想参加这个工程而带来了什么变化。首先，大家都已经同意，如果把一个大脑分成两半，再像我已描述的那样在两个脑半球里输入程序，大脑的整体感觉就能产生的话，那么，如果我们再把各个脑半球仔细分割并且像处理脑半球那样处理这些部分，又何尝不可呢？现在，一个大脑被分成了四个部分，它们不仅被装在单独的瓶子里，而且都有了专门的实验室，这样，就能允许更多的人参加研究工作。很自然，看来没有什么理由能阻止进一步的分割。因而，十个

世纪之后，就出现了这样一种局面：一个大脑被分解为几十亿个神经原，每个实验者都有一个神经原。每一个神经原的两头都装有一个脉冲管，只要输入了接收和发送脉冲的程序，它就能照办。

同时，人们中间也有许多卡桑德的徒子徒孙。他们不再坚持近似条件，因为在大家都希望获得一部分脑组织的情况下，这样做显然会引起公愤。但是，卡桑德的徒子徒孙指出，即便在大脑被分割并散布各处以后，大脑最初的拓扑学结构，即每一个神经原的相对位置和定向态势，还是能够被继续保持的。他们还强调指出，神经原按照相同的时间顺序，也即按照当它们都在大脑中时所展示的那种时间形式通过输入的程序来激活。

但是，关于拓扑学结构的建议总是引来一场冷嘲热讽。人们会提出这样的异议："单个的神经原怎么会知道它同其他神经原相互之间的关系？它是怎样得出这样一个印象的？在感觉产生的一般情况下，神经原为了被激活以产生一定的感觉，要挨次排列，相互处于一定的空间关系，以便真正彼此激活。但是，由于技术的发展，这些原来属于必要的条件不再是必不可少的了。举个例子，如果我们现在要让那位生活在十个世纪以前的小伙子产生感觉的话，这些条件就不再需要了。他大脑中的一个神经原就在我面前的这个瓶子里。如果我们把这些神经原跟一张嘴连在一起，那么，这张嘴就会开口告诉你亲身的感觉。"

至于卡桑德建议中的第二部分内容，读者也许会想，在大脑被一点点分割以后，每部分之间的同步条件已经不需要再加以考虑了，这样，关于单个神经原的激活与其他神经原的激活在时间上的关系，最终会被觉得无关紧要。最初当大脑只被分成两半球时，这个条件也被忽略了。但是，也许会由于忽视了单

个神经原激活的时间和顺序而导致将程序编制技术视作荒谬。这就使得时间和顺序条件现在被悄悄地提起了，只是已没有了卡桑德式的思索。那些人每人守着一个瓶子，里面装有一个神经原，为了使这个神经原产生某种特定的感觉，他们等待着每一个经过恰当程序处理的脉冲传给这个神经原。在这里，神经原激活的“正确”时间顺序只是被假定为产生感觉的某种基本要素。

一晃十个世纪过去了。现在，这个具有几十亿个保养得很好的神经原的世界眼看着就要爆炸了。负责这项工程的是两名思想家。

一个名叫斯鲍勒。一天，他发现自己的那个神经原有点变质了。他和别人一样，又弄来了一个类似的神经原，并把原来的那个扔了。这样，他和其他人一样，也违反了卡桑德的“神经同一性”原则。事实上，这个原则连卡桑德的徒子徒孙们也从没有给予足够的重视。大家认识到，在正常的大脑里，由于细胞的新陈代谢，神经原中的特殊物质一直在更换，而更换结果是形成完全相同的神经原。斯鲍勒的做法无非加快了新陈代谢的速度。除此以外，正如有些卡桑德的徒子徒孙们难以置信地争辩的那样，如果把所有神经原都给更换了，结果对感觉者来说产生了一种新的同一性，那又怎么办呢？每当实现了同样形式的激活，一个感觉者就会仍然获得同样的感觉（甚至连卡桑德的徒子徒孙们也相信，认为这样就等于换了一个感觉主体的说法是讲不通的）。因此，对神经同一性的改变似乎并不会破坏感觉的产生。

这个斯鲍勒，在换了那个变质的神经原以后，照旧耐心地等待着预计原定数个小时后发生的神经原的激活。就在这个时候，突然，他听到玻璃摔碎的声音，紧接着有人大骂。有一个

傻瓜居然把别人装有神经原的瓶子碰翻了，瓶子掉在地上顿时摔得粉碎。那个瓶子的主人这时什么感觉都形成不了了，只有换一只瓶子和一个神经原。斯鲍勒明白，这个可怜的人眼看着就要等到一个脉冲了。

这个人来向斯鲍勒求援了。他说："我以前帮过你，你现在可不能见死不救啊？本来还要五分钟，我就可以等到一个脉冲了，那种感觉就差一个神经原就要被激活了，可现在全给砸了。你能不能把你的那个神经原借给我使一使，我今天真想尝尝激活的滋味！"

斯鲍勒考虑了一下。突然，他眉头一皱，计上心来，问道："你原来的那个神经原是不是和我的一样？"

"是的，完全一样。"

"这就好了。我这个神经原也是刚刚换的。我们有时都这样做。你为什么不把我的神经原连瓶一块端过去放在你原来的位置上呢？这样，因此而产生的感觉同你被摔坏的那个在五分钟以后就产生的感觉完全一样，因为这两个神经原一模一样。瓶子一样不一样完全无足轻重。你用完以后，我还可以拿回来用这个神经原来产生它原定要产生的感觉。请等一等！我们两个人都认为拓朴学的条件纯属胡扯。因此为什么还要把瓶子端来端去？就搁在这儿，你先用，你用完了我再用。我们两个人预期的感觉都会产生。请等一等！这样的话，我们只要激活这一个神经原，而不必都激活所有与它类似的神经原！这样的话，每一类神经原中只需要被激活一个，它反复被激活就会产生所有这些感觉！当神经原被反复激活时，它们怎么会知道只是在重复一个脉冲呢？它们怎么会知道被激活的相对顺序呢？这样的话，在一类神经原中，我们只要使一个神经原被激活一次，因为它完全能够使一切脉冲产生（在从分裂的脑半球到分裂的神经原过

程中，只要我们不考虑同步条件，就能得出这个结论）。这些神经原难道不可以代替任何大脑中自然激活的任何神经原？那么，我们这些人还呆在这儿干什么呢？”

接着，他又产生了一个更加极端的想法。他说：“如果一个神经原的激活就能代表一类神经原的激活，从而引起所有可能的神经感觉，那么，感觉者怎么能相信他是通过获得感觉这一事实，不光同这种最小的物理实在保持着联系呢？我们谈大脑，谈神经原，谈了这么一大通，一切依据在于我们认为这完全基于真正地发现物理实在的基础之上，照现在这种情况，这些说法完全是站不住脚的。也许真的存在一个物理实在的体系，但是，如果这个物理实在通过神经感觉反映到我们的大脑之中，而这种感觉又五花八门，那我们又怎么可能知道究竟什么是对它的真正感觉？因此，对这样一个感觉的信念自我瓦解了。也就是说，除非它经受了卡桑德原则的考验。”

另一位思想家碰巧也姓斯鲍勒。他也得出了同样的结论，但他的推理过程略有不同。他喜欢将神经原排成一串。有一天，他从一串相似的神经原里取出他主管的那个，他记得他应该把这个神经原跟脉冲管联在一起，以使它激活。但他并不想糟蹋整串神经原，因此，就把在这串神经原两端的两个神经原同脉冲管的两极连接起来，并调整好时间，使脉冲穿过整串神经原后准时到达他们的那个神经原。接着，他注意到有一个神经原同一般感觉中的神经原不同，它十分善于同时采取两种激活形式——这串神经原碰巧具有近似性和因果联系，而且还有它曾经激活过的程序感觉。发现这些之后，斯鲍勒开始奚落“神经环境条件”了。他说，“伙计，我能够把这个神经原同你大脑里任何一个神经原连接起来，如果我能够叫它准时激活，我就能够使它畅通无阻地进入这些程序感觉的任何一种，就如同它

盛在我的瓶子里、装在我的脉冲管上那样顺利。”

有一天，问题出现了。一些不属于这项研究工作的人晚上来到实验室，胡乱地摆弄了那些瓶子，使斯鲍勒周围的许多神经原都死了。斯鲍勒怔怔地站在他那个死去的神经原面前发呆，他木然地望着周围发生的一切，他陷入了沉思。这么多的神经原死了，这么多激活的神经原无法得到物理的实现，明天一早，这位感觉者会有什么感觉呢？但是当他环顾左右时，突然注意到了别的东西。现在几乎所有的人都弯着腰在查看自己的瓶子底下被损坏的设备，斯鲍勒突然发现，在每一个瓶子边上都有一颗脑袋，每一个脑袋里有几十亿个各种各样的神经原，每隔一定的时候，大约每一种神经原有几百万次激活。近似条件无足轻重。但是，每当神经原通过瓶子在特定的时间激活时，必要的活动早已在操作者的脑袋里发生了，至少在其中的一颗脑袋里发生了，在这颗脑袋里，一种不太严格的近似条件已获得了满足！每一颗脑袋就是装有神经原的瓶子，就是脉冲管，完全可以实现大脑在物质上的延伸。但是，斯鲍勒转念一想：每一个大脑的每一种感觉都需要有同样的物理实现，因为，所有的大脑都是可延伸的。我的大脑当然也在此例。但是，这样的话，我的思想和感觉就会像浮云，而我的信念正是建筑在这些思想和感觉之上的。它们都是值得怀疑的，连最初使我对这种生理学构造深信不疑的信念也值得怀疑。除非卡桑德的观点是正确的，那么，在一定程度上，生理学也可以被归结为是荒谬的。它自我瓦解了。

这种思想扼杀了这个伟大的工程，连同延伸的大脑。为了探求感觉的本质，人们又转向其他异想天开的活动，去得出新的结论。这又是另一回事了。

D·C·丹尼特

D·R·霍夫施塔特

反　　思

这个离奇古怪的故事乍一看好像是在偷偷摸摸地拆本书其他部分的墙脚，在用归谬法否定大脑与感觉之间既明显又单纯的联系。归谬法如脱轨的火车正滑向它那疯狂的结论，我们怎样才能阻止它呢？下面是几个提示：

假设有人宣称，在他家中有一尊米开朗琪罗的雕塑"大卫"的小型复制品（即使是用大理石制成的），但当你闻讯赶去他家时，你发现立在你面前的是一块20英尺高，大致成长方形的纯白大理石，你不免会觉得迷惑不解。而他会这样对你说："我还没来得及把它拆开，但我知道大卫就在那里。"

再让我们回想一下朱博夫所描述的与大脑各部分相连接的"脉冲管"和"脉冲程序编制器"的情况又是多么贫乏。我们所知道的不过是，这些装置的作用是在一生中，按照恰当的秩序、恰当的时间，给相连接的神经原或一群神经原以恰当的脉冲。我们会认为这纯属无稽之谈。但是，通过考虑实际上是极其"易得"的技术成果，可以反思这些脉冲管事实上产生的东西。大规模的罢工使电视台关闭了，因此也看不到什么电视节目了。国际商用机器公司伸出了援助之手，他们给电视迷们寄去了"脉冲管"。这种脉冲管可以安装在电视机上，里面已输入了程序，能够播出10个频道的新闻、天气预报、短剧、体育等节目。当然，所有的节目都是凭空捏造的（新闻也不是真新闻，但至少看上去

像真的一样）。国际商用机器公司的工作人员说：我们毕竟都知道电视信号只不过是电视台传送的脉冲，我们的脉冲管只不过走捷径帮助电视机收到了信号。那么，这些奇妙的脉冲管里面究竟有什么呢？是某种特殊的录像带？但它们是怎样录制出来的？在没有拍摄真实场面的情况下，如果你想看到逼真的节目，这确实是一项非常困难的工作。如果你要追究来龙去脉，那么，只有真实的世界才能胜任为真实的电视节目提供（并控制）信号来源的重任。创造一个真实的感知世界的重任（笛卡儿在他的《形而上学的沉思》里，把这个任务交给了具有无尽欺骗能力的魔鬼），从原则上讲也许是可能的，但在实际上却是完全不可能的。笛卡儿让他的邪恶魔鬼具有无尽的欺骗能力是正确的——小骗子就会露出马脚，跌回到真实的世界之中，把幻觉变成了现实，尽管现实的形象几经阻滞，几经曲解。

以上这些论点给了朱博夫隐晦的论点以迎头痛击。若将这些论点联合起来，能给朱博夫的论点以致命一击吗？我们可以肯定他的结论是荒谬的，这只要问一问，一个类似论点是否可提出来证明没有必要出版书籍。我们不是可以只要印一次字母表就能代替所有书的出版了吗？谁说要印一个字母表？一个字母或者一笔，一点，不也就够了吗？

逻辑学家雷蒙德·斯莫里安认为，学习弹钢琴的正确方法应该是单独地熟悉每一个音符，一次一个。这样，举例说，你练习中音C也许要花整整一个月的时间。而练习键盘头上的音符只要几天时间就足够了。但别忘了休止符，因为它们在音乐中也是同样重要的。你可以花一天时间来练习全音休止符，花两天时间练习半音休止符，再花四天时间练习四分之一音休止符。一旦你完成了这样一项艰苦的训练，就可以想弹什么就弹什么了！听上去很有道理，但也有点儿不对……

物理学家约翰·阿奇博尔德·惠勒曾想过，所有的电子都十分相似的原因就是因为实际上只存在一个电子，它在时间两端之间不停地来回穿梭，通过无数次重复自己已走过的道路，编织着物质世界这张巨网。巴门尼德的话也许是对的：只存在一件东西。然而，这件东西有其时空的各部分，一些部分与另一些部分在天文位置上结成了许多种关系，这种关系组织在时间与空间上是极其重要的。但是对谁重要呢？对作为观察者的那一部分。可是这如何与这张网的其他部分加以区别呢？

13 我在哪里？

丹尼尔·C·丹尼特

依据“情报交流自由法”，我赢了这场官司。既然官司已经打赢了，我也就能畅所欲言，向大家透露我生活中的一桩奇遇。这桩奇遇，我想不仅会使那些研究心灵哲学、人工智能和神经科学的人觉得趣味无穷，也会使一般的人觉得十分新奇。

几年以前，五角大楼的几位官员来找我，要我自愿接受一项有高度危险的秘密使命。国防部正在同国家航空与航天局以及霍华德·休斯飞机公司联合研制一种超音速钻地装置。按照五角大楼一位官员的说法，此装置是用来以高速钻透地核，携带一颗经过特殊设计的原子弹头，“直接命中共产党的地下导弹基地”。

现在的问题是，他们先前在俄克拉何马州的塔尔萨进行的一次试验中，成功地将弹头射入地下约1英里深的地方，他们要我将弹头找回来。“为什么要我去？”我问他们。唔，这项秘密使命涉及到某些最新大脑研究成果的试用，他们听说我对大脑的研究颇有兴趣，而且也听说我有浮士德般的好奇心，有无所

畏惧的勇气，诸如此类……嗜，我怎么可以拒绝呢？五角大楼的官员所以会登门造访，是因为地下的那颗原子弹头具有极强的放射性。根据放射能检验器的显示，这颗弹头由于本身的原因，再加上与地下深处各种物质发生的化合作用，产生了极强的放射性，它会严重破坏某些脑组织。这种致命的射线对身体内其他组织与器官无害，而要使脑子不受其害却尚未找到办法。因此，他们决定，将去找回弹头的那个人必须把他的大脑留在地面。留在地面上的大脑将得到妥善安置，它通过精密的无线电联系，行使大脑的控制功能。如果我接受了这个使命，就要接受一次手术，我的大脑就要从我的头颅中取出去，然后被安顿在休斯顿航天控制中心的实验室里。他们将在我的大脑和头颅上各连接一个微型无线电收发器，这两个收发器将互通信息，保持密切的联系，半点信息也不会丧失，所有联系一仍其旧。一开始我还有点犹豫。这能行吗？休斯顿的脑科医生却一个劲儿地鼓动我。他们这样对我说："你好好想一想，动这样一个手术，实际上无非是把神经延伸一下。如果你的大脑在你的头颅中移动1英寸，这对你的心灵不会带来任何改变或伤害。我们给你做这个手术，只不过把无线电连接器拼接在你的脑神经上，使它具有无限的弹性。"

他们领着我参观了休斯顿的那个即将贮藏我的大脑的实验室，在那里，我看见一个崭新的容器，如果我同意的话，我的大脑将安放在那里面。他们为这次手术召集了一大批著名的神经病学家、血液学家、生物物理学家，还有电器工程师。他们花了好几天时间进行讨论和研究，最后，我终于同意让他们试一试。手术之前，我接受了一系列的试验，他们先给我验了血，进行了大脑扫描，还进行了实验、面试之类的准备工作。他们详细地写下了我的自述，记下了我的信仰、希望、恐惧和爱好。

他们甚至还记下了我最喜欢的立体声音乐，并临时对我进行了精神分析。

动手术的日子终于到了。当然，手术是在麻醉的情况下进行的，因此，对手术过程我一无所知。当我苏醒过来，睁开双眼，朝四周打量了一番，问了一句人们惯常在手术后难免要问的可悲的话："我在哪里？"护士微笑着俯身回答我："你在休斯顿。"我在想，这句话对我来说又对又不对。她递给我一面镜子。啊，果然不出所料，镜子里的我有一颗奇怪的脑袋。我的脑袋上被凿了一些窟窿，然后他们用钛把一些微型天线固定在窟窿上。

"我想手术很成功，"我说道，"我想看一眼我的大脑。"他们领着我（我有些晕，走起路来有点踉踉跄跄）穿过一条长长的走廊，走进存放着我的大脑的实验室里。那群科学家一看到我，便齐声欢呼起来，我行了一个潇洒的礼，我希望我这个敬礼还像个样子。我被搀扶着来到盛着自己大脑的容器前，隔着玻璃往里瞧。容器里的营养液看上去很像啤酒，我的大脑就飘浮在里面，上面几乎尽是五颜六色的集成线路片、塑料细管、电极以及其他装置。"这难道就是我的大脑吗？"我问道。"你按一下输出收发器的开关，看看会发生什么事情。"项目主任对我说，我把开关拨到"关"，顿时，我觉得天旋地转，恶心阵阵，一头栽在周围人的身上，其中的一位好心地把开关拨到了"开"。我顿时又恢复了平衡。我暗地里思量着："我现在正坐在折叠椅上，透过一块玻璃打量着自己的大脑……但是，等一下，我是不是应该这样想：'我现在正飘浮在营养液里，我自己的眼睛正从外面打量着我'？"我力图想使自己朝后面那个想法想，我想把这个思想透过玻璃传给我那浮在营养液里的大脑，但是不成功。我又试了一次。"我丹尼尔·丹尼特正飘浮在营养液里，我自己的

眼睛正从外面打量着我。"不行，还是不成功。我简直彻底糊涂了。作为一个坚信物理学信念的哲学家，我毫不犹豫地认定我的思想应该发生在我大脑里的某一个部位。但是，每当我想"我在这里"时，这个思想却发生在这里，在容器的外面，在我丹尼特打量我自己大脑的地方。

我试了又试，但都以失败而告终。于是，我又想采用心理练习的办法。我自己这样想："太阳在那里闪耀。"这个思想我连续想了五遍，每一遍心中都是指一个不同的地方，先后顺序是这样的：实验室的朝阳面、医院前的草坪、休斯顿、火星，最后是土星。我完全能够随心所欲，想指哪儿就指哪儿，我可以一会儿指宇宙的尽头，再过一会儿又指我胳膊上某个雀斑的第2象限处。想"那里"的时候竟会这样容易，可为什么一想"这里"，就这么麻烦呢？"在休斯顿这里"没问题，"在实验室这里"也没问题，甚至连"在实验室某一角落这里"也没什么问题，但是，一想"在容器这里"就不行了。当我这样想时，试图闭起双眼，这样好一些，但只是短短的一瞬间，最终还是行不通。我不敢肯定。这个我所不敢肯定的事实还十分烦人。我在想"这里"的时候，我怎么知道这个"这里"究竟指的是哪里？事实上我指的是某一个地方，但我想的却是另一个地方，这可能吗？我不明白，要是一个人与他的那种躲过了大脑科学家、哲学家、物理主义者和行为主义者猛攻的心灵生活之间的密切关系不解除，怎么能够承认这一点呢？看来在当我说"这里"时究竟是指哪里的问题方面，我的成见也许显得根深蒂固了。但是，按照我现在的状况，也许是因为我受了心理习惯的影响，整个关于方位的思想都是错误的，或者是人所在的地方（以及象征语义分析目的的思想所在地）与他的大脑、灵魂的物质位置，不一定是同一个地方。我现在越来越糊涂了，想依靠一个哲学家的策略来确

定自己的位置。我开始给事物取名。

我大声地对自己的大脑说："约里克，你是我的大脑。我身体的其他部分正坐在椅子上，我叫它哈姆雷特。"这样，我们都在这儿了。约里克是我的大脑，哈姆雷特是我的身体，我呢，则是丹尼特。现在，我在哪里？当我想"我在哪里？"时，这个思想究竟是在哪里出现的呢？它是出现在飘浮在容器内的大脑里呢，还是出现在我肩膀上的脑袋里？还是哪儿也不曾出现？它的时间坐标没有什么问题。它是不是也不能有空间坐标？我开始设想各种可能性。

1．有哈姆雷特便有丹尼特。这个原则很容易被哲学家们极感兴趣的大脑移植思维实验所驳倒。如果汤姆和迪克互换大脑，那么，现在的汤姆的身体是迪克的。你问问他。他会说他是汤姆，而且能告诉你关于汤姆过去的最大隐私。很明显，我的身体和我可以分离，但是，我不大可能与我的大脑分开。通过这个思维实验，我们很容易发现这样一种情况，即在大脑移植手术中，谁都想当捐献者，而不愿当接受者。其实我们最好还是把它叫作身体移植手术。也许事实是这样。

2．有约里克便有丹尼特。但是，这种说法一点也不吸引人。我现在明明是在容器的外面，并且还在不好意思地算计着准备回房间去吃午饭呢。既然如此，我又怎么可能在容器里面并且哪儿也不想去呢？我明白，这样说有点强词夺理，但不管怎么样，这已触及重要的问题了。我在为我的这个直觉到处搜罗证据，最后，我忽然想起一个洛克也许会感兴趣的守法主义的论点。

我自己想，假如我就要坐飞机去加利福尼亚州，到那里去抢劫一家银行，然后被警察逮捕了。逮捕之后，我将在哪里受到审判呢？是在发生抢劫案的加利福尼亚州，还是在抢劫犯的

大脑所在地得克萨斯州？我是一个大脑在其他州的加利福尼亚罪犯呢，还是一个遥控同案犯在加利福尼亚作案的得克萨斯罪犯？我差点因为此案的审理权无法确定而逃避刑事责任，当然，也可以认为此案属州际犯罪，即联邦法院有管辖权。就算我被审判并被判了刑，我的身子即哈姆雷特将在加利福尼亚州服刑，而我的大脑约里克却逍遥法外，在得克萨斯州装有营养液的容器内悠然自得，对此加利福尼亚当局会同意吗？得克萨斯州会不会囚禁大脑约里克，而让身体哈姆雷特继续为非作歹？我对这些可能性颇感兴趣。如果得克萨斯州取消了死刑或其他酷刑，那么，他们就要把约里克从休斯顿押送到利文沃斯，在那里，他们必须用一整套生命维持系统来伺候它。我呢，除了名誉受点损害外，其他一点也没有什么，我还是十分自在。如果得克萨斯州当局有兴趣解押囚犯，那么，不管约里克解押到哪里，他们都解押不了我。如果上述情况属实的话，那么，就会产生第三种可能性。

3．丹尼特自己觉得在哪里就在哪里。这种说法归纳起来是这样的：在任何一个特定的时间里，一个人都会有一个观点，观点的所在地（观点在内部是由其内容所决定的），就是那个持观点的人的所在地。

这样一个命题并非没有混乱，但是，对我来说，它朝正确的方向跨进了一步。不过，这个命题在所在地的问题上有一个不足之处，因为，这种说法就同掷硬币打赌一样，正面我赢，反面你赢，无论怎样，总不会错。关于我在哪里这个问题，我不也经常弄错至少是常常不能肯定吗？你难道就不会迷路吗？当然，除了在地理上迷路之外，在别的意义上你也会迷路。如果你在莽莽的大森林里迷了路，你还可以这样安慰你自己，你至少还知道你在哪里，因为，你就在你身体熟悉的环境这里。也许，在

这种情况下，你不必多责怪自己。尽管如此，天有不测风云。也许会有更糟糕的困境，我也不敢肯定我现在就没处在这样一个困境里。

显然，观点同个人的所在地确有关系，但它本身并不十分明确。一个人观点的内容既不同于信仰或思想的内容，也不受它们的制约。比方说，在看立体电影时，银幕上一列火车迎面驶来，观众们还以为身临其境，吓得在座椅上尖声喊叫、辗转不安。我们怎样解释这些观众的观点？他们难道忘了此刻正安全地坐在电影院里吗？这里，我想说明这样一个问题，在观看立体电影时，观众感觉到了观点的幻移。在其他场合，这种幻移不那么强烈。在实验室和工厂里从事危险操作而使用反馈控制的机器臂和机器手的工人也经历了观点的幻移，这种幻移比起立体电影产生的幻移效果更加逼真、更加明显。通过机器手，他们能够感觉到他们正在搬运的集装箱又滑又重。他们对自己在哪里这一点再清楚不过了，他们决不会因感觉而产生错误的信念，但是，他们又好像置身于他们探眼观望的隔离室之中。由于心理作用，他们能够来回转移观点，就好像在你眼前画一个透明的内克尔氏立方体，或是作一幅埃歇尔的画。如果我们认为通过做这样一点心理体操，他们就能来回转移他们自己的话，那就有点过分了。

尽管如此，这个例子还是给了我希望。如果我确实是在容器内，那么，我也会习以为常，慢慢熟悉这种观点。我该专心致志地认为，自己悠然自得地飘浮在营养液内，把我的意志发送给在外面的身体。我认为，不管这项任务是困难还是容易，都与大脑的所在地无关。如果在这次手术之前我已经在这样做的话，那么，我现在就会发现这是第二天性。你自己不妨也试一试这种假象。让我们设想这样一种情况。你写了一封富有煽动

性的信，这封信发表在《纽约时报》上。结果，政府下令监禁你的大脑，把它关押在位于马里兰州贝塞斯达的"危险大脑诊所"里，监禁期为三年。当然，你的身体还是自由的，可以挣一份工钱，攒起来让政府征税。而就在此时此刻，你的身体正坐在一个礼堂里，倾听丹尼尔·丹尼特给你讲他自己相似的遭遇。你不妨试一试。你就想你自己在贝塞斯达，然后返回来追寻你的身体，虽然很遥远，但看上去却又很近。正是由于这种远距离的监禁(是你自己的，还是当局的?)你才得以控制自己的脉冲，使自己的两只手能在演讲结束之后彬彬有礼地鼓几声掌，然后驾驭你自己的身体，到休息室去喝一杯理应享用的雪利酒。想象的任务当然是艰巨的，但是一旦达到目的，结果将是令人满意的。

不管怎么样，我还在休斯顿，可以说沉浸在冥思遐想之中。但是不久，我的思绪很快就被休斯顿的医生们打断了，他们在我动身去执行那项危险的使命之前，想检测一下我那新修复的神经系统。我前面已经说过，开始时我感到有点晕眩，这毫不奇怪，但不久我就习惯了我的新环境（实际上我的新环境和旧环境并没有什么区别)。然而，我的调节系统却不太理想，至今还有一些小小的协调障碍存在。光速虽快，但却有限，而当我的大脑和身体分离越来越远时，我的反馈系统的微妙的相互作用却因时滞而出现了混乱。如果你不能同步听到自己的说话声，你就几乎无法正常说话。同样道理，由于我的大脑和身体相隔数英里之遥，我的眼睛几乎无法跟踪一个移动的目标。我现在已无法像以往那样击中一个呈曲线慢速移动着的球，但是，这个毛病在多数情况下还发现不了。当然，对此也有若干补偿。我对液体的味觉一如既往，吃酒时仍觉得喉咙发热，只是肝脏受损，不过我的酒量明显见长，几乎到了百杯不醉的程度。对于

这一点我的一些挚友已有所察觉(尽管我有时也佯装喝醉，以免使大家对我百杯不醉的海量引起注意)。出于同样考虑，我的胳膊扭伤以后，我也口服一些阿斯匹林，但是，如果疼痛难忍，我就告诉休斯顿在体外给我注射些镇痛的可待因。因此，每当我一生病，光是与休斯顿联系的电话费就不得了。

现在还是言归正传，讲讲我的地下历险吧。现在，百事齐备，医生们和我都十分满意。于是我整装出发，把大脑留在休斯顿，自己登上了飞往土尔萨的直升飞机。不管你会说什么，在我看来就是这么一门事。我就这么讲，我仿佛有点不由自主了。一路上，我思索着手术前的种种担忧，并且确定我在手术后最初的思考包含着混乱。事情并不像我以前想的那样离奇和玄乎。我在哪里？很明显，分开在两个地方。正如有人可以一只脚踩在康涅狄克州，而另一只脚踩在罗得岛州一样，我也是同时在两个地方，既在容器里又在容器外。我们以前常听到关于四分五裂的人的故事，我实际上就是这样一个人，对于这个答案，我越想越对。但是，奇怪的是，答案越显得对，这个正确答案所针对的原来的那个问题却越显得不重要，古往今来，一切哲学问题都会遭受这种悲惨的命运。当然这个答案并没有完全使我满意。问题还有，而且我也想求得它的答案。这个问题不是“我的各个部分在哪里?”也不是“我目前的观点是什么?”或至少说似乎有这样一个问题。因为，从某种意义上说，钻到地下去寻找那颗原子弹头的是我，而不是我的大部分，这个事实看起来确实是无法否认的。

当我找到原子弹头时，我随身携带的特制的盖格计数器上的指针不灵了，我这时候真庆幸自己把大脑留在地面上了。我于是用普通无线电与休斯顿通话，把我的位置和进展情况告诉那里的控制中心。他们回话要我拆毁弹头，并根据我的现场观

察，给我拆毁弹头的行动指令。我便用切割机开始拆除弹头。就在这时，可怕的事情发生了。我突然什么都听不见了。起先，我还以为是我的无线电耳机坏了，我敲敲头盔，一点声音也没有。很明显，我那助听的无线电收发机出了故障。我再也听不见来自休斯顿的声音了，也听不见自己的声音了，但是我还能讲话，于是我就向休斯顿报告了这里发生的一切情况。正说着，另外一件可怕的事情又发生了。我的发声器官麻痹了。紧接着，我的右手也出了毛病——另一个无线电收发机坏了。我真是倒了大霉。这还不够，接踵而来的事情更为可怕。没过几分钟，我的眼睛瞎了。我先是怪自己的运气不好，再是怪休斯顿的那一大帮科学家把我弄到这么危险的地步。我又聋又瞎又木，呆在土尔萨地下1英里多深的放射性的洞里。接着，我大脑中最后一点无线电联络也中断了，我突然身陷绝境，面临着一个更新更惊人的问题：就在我马上要被活埋在俄克拉何马州地下之前的一瞬间，我已在休斯顿脱离了我的身体。我并没有马上意识到自己的新身份。令人焦虑的几分钟过去了，我终于醒悟过来，我的可怜的躯壳此刻正在几百英里以外的地方，它的心还在跳动，肺还在呼吸，但却与死人无异，就像一个心脏捐献者的躯体，我那颗脑袋里塞满了已经出了故障的电子元件。最初我曾觉得观点的转移似乎令人不可思议，但现在却觉得十分自然。尽管我能在土尔萨的地下想着回到我的躯壳里，但要保持这个念头却不很容易。因为，如果还以为我仍在俄克拉何马州的土尔萨，无疑是一个错觉：我与那个躯体现在已毫无联系。

脑中带着某种应对之表示怀疑的启示，我想到，我出于偶然，发现了基于物理主义原则和前提之上的灵魂非物质性的有力证据。因为，随着土尔萨和休斯顿之间最后一个无线电信号的消失，我难道没有以光速从土尔萨飞到了休斯顿？我难道没

有在不增加任何质量的情况下，完成了这一过程？以如此之神速，从A地飞到B地的肯定就是我自己，或者说是我的灵魂或心灵，这是我存在的无质量的中心，也是我的意识的故土。我的观点多少有点落后了，但是，我已经注意到观点对于个人位置的间接影响。我想一个物理主义哲学家对此不会再有什么可说的了吧，除非他采取不许谈论个人的极端而反直观的做法。但是，个人这个观念在每一个人的世界观内已经根深蒂固，以致在我看来任何否定这个观念的做法都会像笛卡儿"非存在"这个否定命题一样，显得极其无力而虚假。

我当时身临绝境，心灰意懒，哲学发现给我带来了喜悦，帮我熬过了这段难忍的时光。我全身疼痛、恶心，由于躯体不在，不知这些感觉从何而来，因此使得这一切更加可怕。胳膊上没有肾上腺素大量分泌而带来的刺痛，没有怦怦的心跳，也没有先兆性的唾液分泌。我确实感到腹部某处有剧痛，于是，我便错误地希望，我现在正在经受一场正好与肉体解脱相反的转变。但是，这次腹部剧痛与以往不同，而且与其他部位也没有什么联系，于是，我便认定这不过是一个幽灵身体的第一个幻觉，同别的被截肢者一样，我肯定也要经受这些幻觉的折磨。

我心乱如麻。一方面，我为自己的哲学发现而欣喜万分，而且正挖空心思（这是我还能做的熟悉事之一）想把我的发现写成文章，在刊物上发表；而另一方面，我又痛苦，孤独，充满了恐惧与不安。幸运的是，痛苦与孤独并不长久，我的那个专家小组设法使我安睡，暂时忘掉这一切。等我从甜蜜的睡梦中醒来时，我听见了优美的音乐声。啊，正是它，正是我最喜爱的勃拉姆斯钢琴三重奏开始时的旋律。他们曾把我喜爱的乐曲一作了记录，原来如此！很快我就发现，我听音乐并没有用耳朵，从唱针尖下传出来的音乐，通过某些奇怪的整流电路，直

接输入了我的听觉神经。勃拉姆斯的音乐就像输液一样注射进了我的体内。对于任何一个立体声爱好者来说，这种经历是难以忘怀的。一曲终了时传来项目主任那令人温暖的声音，对此，我一点也不感到吃惊。他是对着扩音机说话的，现在扩音机就是我的耳朵。他证实了我对各种故障的分析，并向我保证，专家们正在采取步骤，使我重新恢复躯体。他的讲话简单明了。在这之后，我又听了几首曲子，便昏昏沉沉地睡着了。后来我才知道，这一觉就是大半年，我醒来时发现，我的各种感官完全恢复了正常。我照了照镜子，不禁大吃一惊，我已变得面目全非。我蓄了胡子，太浓了一些，看上去还是那样顽皮、聪明、性格刚毅，无疑与原来那张脸有许多相似之处，但这肯定是一副新颜。我又对自己的一些功能作了验证，结果，我毫不怀疑地认为这是一个崭新的躯体。项目主任也证实了我的结论。他对我现今躯体的过去经历只字未提，我也不想弄个水落石出（后来回想起来，我这样做很明智）。许多哲学家对我的困境不了解。他们最近认为，新躯体的获得对人不会有任何改变。经过了一段时间对新声音、新臂力等情况的适应，总的说来，我原来的人格还是得到了保留。做过整形外科手术的，都会发生戏剧性的变化，更不必说做过性别改变手术的人的人格。我想在这种情况下，没有人会认为人格还能完整地保留下来。不管怎么样，我对我的新躯体很快就适应如常了，我自己的意识或记忆一点也不觉得它是一个陌生人。镜中的形象也完全习惯了。顺便说一句，我自己头上还装着小天线，因此，我心里很清楚，我的大脑还被保留在它的庇护所里，也就是说，在休斯顿实验室的生命维持系统内。

我决定去探望好心的老约里克。我们可以把我的新躯体叫做福廷布拉斯，它和我一起走进了熟悉的实验室，在那里的专

家鼓掌欢迎我们的到来，当然，他们祝贺的并不是我，而是他们自己。我又一次站在那个容器前，端详着可怜的约里克，我一时心血来潮，带着骑士般的风度关掉了输出发送机的开关。真没想到，居然什么事也没有发生，我感到万分惊奇。我既不恶心，也不头晕，也没有别的症状。一个技术员赶快跑去把开关打开，可我还是没有什么感觉。我要他们给我解释一下这是怎么一回事，项目主任赶紧向我说明了情况。早在给我动手术之前，他们就为我的大脑制造了一个计算机复制品，这是一个巨大的计算机程序，里面储有我的大脑的完整的信息处理系统和计算速度。手术后，他们没有敢马上就送我去土尔萨，在这段时间里，他们同时操纵约里克和那个计算机系统。来自哈姆雷特的输入信号在送给约里克的收发报机的同时，也送给了机算机程序。来自约里克的输出信号不仅发送给了我的躯体哈姆雷特，同时也被录制下来，并同来自程序的输出进行了核对。这个程序取名为“休伯特”，我也搞不清楚为什么要叫这样一个名字。过了数天甚至数个星期，两种输出都同量，并且完全一样。当然这并不能证明，它们成功地复制了大脑的功能系统，但经验说明结果是令人鼓舞的。

休伯特的输入以及由此而产生的活动，在我从肉体解脱出来的那些日子里，同约里克是平行的。现在，为了证明这一点，他们实际上早已让休伯特控制了我的身体。当然这身体不是哈姆雷特，而是福廷布拉斯（我得知，哈姆雷特再没有从土尔萨的地下出来，现在也许早已化成灰烬了。在我的坟头上，还放着那个已被遗弃了的试验品，上面醒目地刻着“超音速钻地装置”的字样。100年以后，考古学家们会对这样的古遗址感到惊奇不已，他们会觉得先人的葬礼颇为奇特）。

实验室的工作人员向我介绍了控制开关。这个开关上有两

处标记着B和H，B代表大脑（他们不知道我的大脑名叫约里克），H代表休伯特。开关现在在H的位置上，他们向我解释说，如果我愿意，可以把它拨回到B的位置上。在我动手拨开关的时候，我的心提到了嗓子眼里（我的大脑仍装在容器里）。结果，只听到"咔嗒"一声，其他什么也没有发生。现在开关被拨在B上了，为了检验一下他们的话是否正确，我又拨开了接通容器的约里克的输出发送机的开关，嗨，真灵，我开始觉得头晕目眩了。等到开关被拨回去后，我的神志又恢复了正常。于是，我便来回拨动控制开关。我发现除了"咔嗒"、"咔嗒"的声响外，没有任何别的变化。当我一句话讲到一半的时候，拨动了开关，这句话的前半部分是在约里克的控制下说出来的，而后半部分是在休伯特的控制下说完的，中间没有丝毫中断的迹象。我有一个备用大脑，如果日后约里克遭到不幸，它会很好地替代约里克。或者，我也可以先用休伯特，把约里克作为备用。不管我先用哪一个，都不会有什么区别，这是因为，大脑长在肩膀上也好，躲在生命维持系统内进行遥控也好，肉体的损耗对它都不会有丝毫的影响。

不久，我就领悟到，这个新发明只有一点令人不安。有人可以将备用大脑休伯特或约里克与福廷布拉斯分离，而将它安在某位后来的名叫罗森克兰茨或吉尔登斯特恩的人身上，那么，很明显，就将出现两个人，一个是我，另一个则是我的超级孪生兄弟。如果出现两个躯体的话，一个受休伯特的控制，而另一个则受约里克的控制，那么，究竟哪一个算是真正的丹尼特呢？不管别人怎样认为，哪一个是我呢？就因为约里克与丹尼特原先的身体哈姆雷特曾经关系密切，我就应该是大脑叫约里克的哪个人吗？这未免有点太守法主义了，太容易使人联想到血缘关系和合法财产的任意性了，在形而上学的层次上，是不

能令人信服的。这是因为，如果在第二个身体出现之前，我一直把约里克作为备用大脑，而让休伯特控制我的身体福廷布拉斯，那么，休伯特与福廷布拉斯这一对，依据法律（一种法律直觉与另一种法律直觉相对立）就是真正的丹尼特，是丹尼特一切财产的合法继承人。当然，这是一个很有意思的问题，但另外还有一个问题，我觉得更为迫切。我有一个强烈的直觉，只要这两对中有一对安然无恙，那我就还活着，但是，对于我是否应该要求这两对都活着，我的感情就很复杂了。

我将我的烦恼告诉了专家们和那个项目主任。我向他们解释说，一想到有两个丹尼特活在世上，我就感到十分惊恐，这多半是由于社会的原因。我不愿意自己成为妻子的情敌，我不愿意有两个丹尼特分享我一个教授的菲薄的薪水。更令人心惊肉跳、坐卧不安的是，你竟会对另一个人知道得那么多，而那个人对你的隐私也了如指掌。我们怎么敢正视对方呢？实验室里的专家们争辩说，我没有看到事物光明的一面。你不是有许多事情要做，而你一个人忙不过来吗？一个丹尼特可以呆在家里做学问，享受天伦之乐，而另一个呢，则可以浪迹天涯，四处闯荡，当然，这个丹尼特也会思念家室，但他一想到另一个丹尼特会把家里安排得井井有条，心里也就无所牵挂了。你既能对爱情忠贞不渝，也能放荡不羁。你甚至还能与自己的妻子浪漫一番。不仅如此，我的同事们七嘴八舌替我设想了种种更为放肆的可能，我的想象力有限，无法在此一一枚举。但我在俄克拉何马州（还是在休斯顿？）所遭受的痛苦使得我不再富于冒险精神，对于这种机会我不敢贸然接受（事实上，我从来就不知道我最初是否真的有过这种机会）。

另外一种前景更加不堪设想：那个备用大脑不管是约里克，还是休伯特，同福廷布拉斯的输入相脱离而独立存在。如果这

样，同前一种情况一样，也会出现两个丹尼特，或者说，有两个要求索取我的姓名和财产的人，一个带着叫福廷布拉斯的身体，而另一个可怜得很，连一个身体都没有。不管是自私还是利他主义的倾向都让我阻止这种情况的发生。所以，我就要求采取措施，不经我的(我们的？不，是我的)容许，任何人不得擅自触动无线电收发机的通信线或者控制开关。我自然不会为了看护设备而把生命耗费在休斯顿。因此，双方一致同意，实验室里的所有的电子通信线都要仔细锁好。控制约里克的生命维持系统的设备以及控制为休伯特提供能源的设备，都配备了具有自动防止故障性能的装置，我随身携带着唯一的控制开关，不管走到哪儿，都能用无线电来遥控那里的一举一动。我把这开关别在腰间，等一等，你瞧，就在这里。每隔几个月，我都要拨一拨开关，看看是否一切正常。我拨开关的时候，自然要有朋友在场，不然的话，一旦开关的另一头无信号，或者被占线，这种事万万不可发生，我将在虚空中永远死去，所以，我需要贴心的人，当出现问题时，再把开关替我拨回来。这是因为，由于这一拨，尽管我还有感觉、视觉和听觉以及任何可随之产生的感觉，但我已无力控制它了。顺便提一句，开关上故意没有标明记号，这样，我永远也不会知道我是从休伯特转到约里克呢，还是从约里克转到了休伯特（你们中的一些人也许会想，在这种情况下，我根本不可能知道我是谁，更不用提我在哪里这个问题了。但是，这些顾虑并不会影响我的丹尼特本质，也不会影响我自己意义上的我是谁这样一个问题。如果从某种意义上说，我并不知道我是谁这种说法属实的话，那又成为另一条具有重大意义的哲学真理了）。

总之，至今每当我拨动开关，都没有发生过什么。那么，就让我们来试一试吧……

“谢天谢地！我还以为你永远也不会拨那个开关呢！你简直无法想象这两个星期我是怎样熬过来的，现在该轮到你下炼狱了，我多么渴望这一刻！你看，差不多有两个星期了——对不起，女士们，先生们，但我必须向我的……唔……我的兄弟解释一下，我想你们会以为他是我的兄弟，但他刚才已经把事实向你们作了解释，因此，你们会明白差不多在两个星期以前，两个大脑开始有点不同步。我现在跟你一样，不知道现在我的大脑究竟是休伯特呢，还是约里克，但是，不管怎样，这两个大脑分离了，当然，一旦分离，它们之间的差距就会越来越大。因为，我处于稍稍不同的接受状态之中来接收我俩收到的输入，这小小的差异很快就加大了。我以为我还控制着我的身体，即我们的身体，这个错觉从来就不曾消失过。我什么也没法做，我无法叫你。因为你甚至不知道有我存在！这就好比被关在一只囚笼里，更确切地说，就好像我被别人占有了，听到我自己在说不愿意说的话，无可奈何地看着自己的双手在干不愿意干的事。你要是在身上搔一把，就会搔动我们两个人的痒处，但不是按我愿意的那种方式，你整天来回折腾，我根本没有办法合眼。整天这样，谁受得了？你整天都忙个不停，我都快给累垮了，我之所以能够坚持到现在，完全靠着心中的信念，我相信总有一天你会拨动开关。

“现在该轮到你了，可你比我要好，因为，至少我还知道你在那里面。我就像一位怀孕的母亲，我吃——或尝、闻、看，现在都是为了两个人，我会十分注意，尽量使你舒服。你别着急。等学术讨论会一结束，我们俩一起飞回休斯顿，再想办法弄个身体。你如果想要个女身也可以，肤色也全由你自己决定。但先让我们斟酌一下。如果我们两个人都想要原来这个身体，为了公平起见，我就让项目主任抛一枚硬币，那我们两个就可凭

运气，看谁能保留这个身体，谁能得到那个新的身体。这样做很公正，对不对？不管怎么样，我会照顾好你的，这一点我可以向你保证。在场的都是证人。

“女士们，先生们，我们刚才听到的这一段话并不完全是我的本意，但你们尽管放心，他说的一切都是实话。现在，如果你们不介意的话，我想我，不，我们最好还是坐下来。”

D·C·丹尼特

反　　思

你刚才读到的这一段故事，不仅不是真的，而且也不可能是真的(你不必猜疑)。故事里那些先进的科学技术根本不可能实现，有的永远也不会实现。但这并不是问题的关键。问题的关键在于，整个故事当中，是不是有些什么东西从原则上讲是不可能的，是前后不一致的。如果哲学幻想变得太离奇了，比如说时间机器啦，双重宇宙性啦，具有无限欺骗能力的魔鬼啦，诸如此类，那么，我们就应该明智地意识到，从中不可能获得什么。我们知道，其中涉及的一些问题也许是不可靠的。是由生动的想象制造的幻觉。

在这个故事里所描述的那种大脑移植手术和微型无线电，在现阶段是完全不可能的，甚至在能够预见到的将来也是十分渺茫的，但毫无疑问，它们是“没有害处的”科学幻想。至于丹尼特大脑约里克的计算机复制品休伯特是不是有可能出现，那就更难说了(作为科学幻想的鼓吹者，我们当然可以随时设想出种种规则，但同时又痛苦地意识到，我们所讲的故事根本

没有理论意义）。我们设想，休伯特和约里克几年来一直处于完全同步的状态，它们之间没有发生过任何联系。这不仅是伟大的科学成果，而且也近乎奇迹了。如果使计算机能差不多像人脑那样，同时快速地处理几百万输入和输出的信息，那就不光要使这样一台计算机的基本结构与现阶段的计算机完全不同。即使是我们制造出了类似人脑的计算机，其体积和复杂程度也必定十分惊人，结果定会使独立的同步行为成为空谈。如果在两个系统内没有同步和同样的处理过程，那么，我们势必要将故事中的一个基本论点抛弃。为什么？因为一个人有两个大脑（其中一个是备用的）的前提是基于这个特征之上的。罗纳德·德苏泽对类似的情况作过以下评述：

> 杰克尔博士变成海德先生时，发生了一件奇特神秘的事情。是不是两个人在一个身体内轮流执政？还有一件事情比这更为奇特：米格尔博士和博格尔博士也在一个身体内轮流当家。但他们就像一对一模一样的孪生子！你犹豫了：为什么不能认为他们已变成了对方？为什么不能，这是因为：如果杰克尔博士能够变成另一个人海德先生，那么，米格尔要变成博格尔就更容易了，因为博格尔跟米格尔本来就长得一模一样。
>
> 我们生来就认为，一个身体至多只有一个主体；我们需要激烈的辩论来动摇这种信念。
>
> ——选自“理性的侏儒”

“我在哪里？”这篇论文中几个重要的论点都基于这样一个假设之上，即休伯特和约里克之间存在着独立的同步信息处理

过程。既然如此，我们应该注意到，这个假设是十分粗暴的。它相当于这样一个假设，在宇宙的某一角落里，存在着一个与地球一模一样的星球，在这个星球上，有你，有你所有的朋友，也有地球上的一草一木，简直是丝毫不差。也相当于这样一个假设，即宇宙只有5天的历史（当上帝在5天前创造宇宙的时候，他创造了许多顷刻间便装满“记忆”的成年人，创造了藏满古籍的图书馆，还创造了许多布满崭新化石的山脉，既然是这样，5天的时间未免也太短了）。

像休伯特这样的人工复制大脑，从原则上讲只是一种假设而已，当然，一些略为逊色的人工神经系统也许在不久的将来就要出现。为盲人制造各种粗糙的电视眼睛已经出现了，其中有些电视眼睛能把输入信号直接送给大脑的视觉皮质，有的则避免了这种精密的手术，而是通过其他外部感受器把信息传递过去，比方说，通过装在手指尖上的触觉感受器，或者在感觉主体的前额、腹部或背部布上一些刺激点，以产生视觉的效果。

这种非手术性的心灵延伸的前景将在下一篇选文中作进一步的探讨，这篇选文是“我在哪里？”的续篇，作者是杜克大学的哲学家大卫·桑福德。

14 我曾在哪里?

大卫·霍利·桑福德

丹尼尔·丹尼特，更准确地说，是丹尼特团体的一个代表，在查珀尔希尔的一个学术讨论会上作了一篇题为“我在哪里?”的讲演，受到了与会者前所未有的热烈称赞。我由于当时正在度假，所以没能和那里的哲学家们一起鼓掌叫好。尽管我的同事们都还以为我住在纽约，潜心于我的研究，但实际上，我正在国防部完成与丹尼特团体密切相关的一项秘密使命。

丹尼特这个人对自己的本质、整体和同一性问题那样关注，以致忘记了这次使命的主要目的是为了找回埋在土尔萨地下具有强烈放射性的原子弹头，而不是为了让心灵哲学中的一些原来已经难以捉摸的问题变得更加难以捉摸。丹尼特告诉我们，他那受遥控的无脑身体哈姆雷特刚开始寻找那颗原子弹头，与无躯体的大脑约里克之间的联系就中断了。他认为哈姆雷特不久就化为灰烬了，对于那颗原子弹头就既不了解也不关心了。我呢，碰巧在最终找回弹头这件事中起了关键的作用。尽管我起的作用与丹尼特差不多，但其中仍存在着一些重要的差别。

丹尼特，或者说是约里克，独自呆在生命维持系统内，与活人体失去了任何直接的或遥远的联系，在这段时间里，被注射了一点勃拉姆斯的音乐。从唱针下传出来的经过特殊处理的音乐直接输入了听觉神经。有些科学家或者哲学家可能会问："既然我们可以不通过中耳和内耳就把音乐输入听觉神经，那么，我们为什么就不可以也不通过中耳和内耳把其他任何信息都直接输入听觉神经呢？既然可以这样做，那我们为什么就不可以也同样绕过这一结构，把信息直接输入大脑中的信息处理系统呢？或者直接输入更深的结构呢？"有些理论家，也许不会是丹尼特，想了解这种以人工信息处理手段代替自然的信息处理系统的过程何时才能达到听觉体验的最终拥有者，人的真正核心，心灵的真正所在地。也有人宁可把这个过程看作是由表及里逐层的转换，即从有机的意识主体逐步转变为人工智能。但是，给约里克的听觉神经直接注射勃拉姆斯钢琴三重奏乐曲的那个科学家却另有自己的看法。他想知道，为什么要费心思把丹尼特的耳朵同其听觉神经分离开来呢？他觉得，如果容器内的大脑以正常方式与耳朵相连，耳朵上套有耳机，然后再在正在土尔萨地下探险的身体上安装一个麦克风，这样做会有许多好处。认为放射线只能破坏脑组织的想法是完全错误的。事实上，哈姆雷特的内耳会首先遭到放射线的破坏，紧接着，身体的其他部分也会被杀死。如果用麦克风来代替哈姆雷特的内耳，而把耳朵留在约里克那里，再在耳朵上戴上一副耳机，这样的话，丹尼特听立体声音乐的效果比起仿照通常的立体声录音而将立体声磁带的乐曲注入耳神经的做法要好得多。如果哈姆雷特在音乐厅里听一场音乐会，他的头稍有移动，远在休斯顿的耳机的输出就会出现细小的变化。这样做就能保持两个信号之间音量和时滞方面的一切细小的差别，尽管这些差别并不能被意识所察

觉，但对于确定音源的方位却至关重要。

上面我描述了如何改进耳机的方法，这种描述也很好地解释了国家航空与航天局的技术人员更为激进的做法。从丹尼特的历险里，他们发现肉眼无法长时间经受来自埋在地下的弹头的强烈辐射。因此，我们应该把丹尼特的肉眼同约里克一起留在地面上，而在哈姆雷特的眼窝里装上微型电视照相机。在我即将赴地下去完成这项秘密使命之际，技术人员已经大大改进了这种眼睛录像机。这种录像机能看，就像耳机能听一样。它们不仅能够把图像投影到视网膜上，而且还能够监视眼珠的每一次转动。每出现一次快速眼动，相机也会迅速作出相应反应；头部的每一次移动都会立刻引起相机位置的相应变化，等等。在多数场合中，用不用眼睛录像机，无多大区别。只是当细看极小的字体时，我才会觉得这种录像机不够敏锐；还有，如果对该系统作精密的调整之后，眼睛录像机在夜间倒看得更清晰。

最为惊人的模拟装置是人工触觉。它对皮肤和皮下的感觉就同耳机对听觉的作用一样，但在描述人工触觉之前，我先要谈谈一些可由眼睛录像机来完成的实验。有关晶状体倒相的实验很容易，只要把相机倒置就行了。同样普通类型的种种新实验也并不难，只要把相机放置在另一些偏离正常位置的方位就可以了。这里有一些例子：所谓兔子放置法，即把两个相机放在相对的而不是并排的位置上；配有广角镜的兔子放置法，其视界可达360°；还有所谓银行或超级市场放置法，即将两个相机放在主体所在的房间里相对的墙上。这种放置法对我们来说很不习惯，需要慢慢熟悉。顺便说一句，用这种放置法可以同时看见一个透明立方体所有的边。

可你们最感兴趣的还是触觉装置。这是一种质量很轻的、

多孔的物质，将它紧贴在皮肤上，可以延伸人的触觉范围，就像无线电和电视能够延伸人的听觉和视觉范围一样。当一只装有触觉传感器的假手抚摸一只湿漉漉的幼犬时，包在触觉接受器内的真手上的皮肤神经就会像真手在抚摸那只幼犬时一样受到刺激，当触觉传感器触摸到温暖的物体时，包在接受器内的真手的皮肤虽然没有因此而发热，但相应的感觉神经却好像真遇到热物一样被刺激起来了。

为了找回那颗深埋在地下的弹头，人们派遣了一名机器人去完成这项使命。这个机器人身上没有活细胞，规模与人体差不多，外表装有触觉传感器，头部装有话筒和照相机，可以向耳机和眼睛录像机传送信息，它的四肢关节与我的差不多，活动起来也与我差不多灵巧。它没有嘴、额以及其他用于呼吸和消化的器官。在嘴的那个部位有一个喇叭，它可以把来自我嘴前的那个话筒接收的所有声音都播送出去。

除此以外，在我与那个机器人之间还有另一个奇妙的通信系统叫做运动和阻抗系统。这种运动阻抗系统薄膜罩在人体触觉表皮上，如果是机器人，这种薄膜就置于触觉传感器的下面。这种薄膜的详细情况我不甚了解，但要说出它的功能却不难。它使得机器人能够准确、同步地模拟人体的大部分动作，同时，机器人四肢所遇到的压力和阻抗也都可以传递到相对应的人的四肢上面。

国家航空与航天局的专家们并没有把我象丹尼特那样一分为二，而是保持我的整体。我将作为完整的人继续留在休斯顿，控制着完成地下使命的机器人，但不遭受辐射。专家们认为，我跟丹尼特不一样，我不会像他那样因为冥思苦想那些关于自己位置的深奥哲学问题，而在完成原来的任务上分心。他们一点也不了解我。

丹尼特曾经提到使用反馈控制的机械臂和机械手来处理危险物品的实验室工作人员。我跟这些人差不多，唯一的不同是我使用的是一个由反馈控制的完整的人，它具有人造听觉、视觉和感觉。尽管看上去我好像就在土尔萨的地下作业，实际上，我很清楚自己究竟曾经在哪里，我正舒舒服服地坐在实验室里，戴着耳机、眼睛录像机，装着人工触觉传感器和运动阻抗系统，对着话筒说话。

但是，事实证明，一旦等我穿戴整齐，我就情不自禁地从机器人的地位来判定自己了。丹尼特首脑分家以后，非常想看一眼自己的大脑，我也是这样，我非常想看看自己穿戴好电子装后会是什么模样。正如丹尼特很难将自己与自己的脑子看作一回事一样，我也很难将自己与那个每当机器人的头移动，他的头便移动，每当机器人的腿在实验室里转悠，他的腿也一起走动的身体看作是一个人。

我也模仿丹尼特给这些东西起名字。丹尼特用“丹尼特”，我呢，用“桑福德”。这样，“我在哪里？”和“桑福德在哪里？”这两个问题答案是一样的。我的名字“大卫”代表那个被保存在休斯顿的由碳水化合物组成的身体，我的全称的中间那个名字“霍利”暂时作为机器人的名字。

很明显，霍利在哪里，桑福德就在哪里这个总的原则是错误的。此刻，那个大卫动一动，它也跟着动一动的机器人正陈列在一个分类精[illegible]的科学博物馆里，而桑福德却没在那里。

还有，机器人在被大卫控制之前或之后，也可以由其他有血有肉的身体来控制。如果说霍利在哪里，桑福德就在哪里的话，那么，只有当霍利与大卫或者大卫的复制品保持着至少像我们前面已经描述过的那几种联络关系时，才能这样说。因此，丹尼特的第一个原则，哈姆雷特在哪里，我就在哪里，需要加

以类似的限定。

当我发现机器人不止一个时，我的那个想把机器人命名为“霍利”的做法碰上了麻烦。在休斯顿，他们已制作了两个与真人差不多大小的机器人，其中一个用的材料主要是塑料，另一个的主体部分用的是金属。从外表看，它们俩几乎一模一样，而且，如果你明白我说这话的意思，它们内部的感觉也是一样的。这两个机器人都没有被送往土尔萨。他们还另外制作了一个机器人，它的尺寸只有真人的3/5，这样，它在窄小的工作舱里可以灵活一点。正是这个机器人找回了弹头。

一旦我面临着不止一个机器人这样一个事实时，技术人员就往往没等大卫睡着便转换了控制他的替身的开关。当那个小霍利从土尔萨载誉凯旋归来之后，我们三个，或者说三个我，便开始轮流执政。当然，这是在三个技术人员的帮助下进行的，以防止机器人因暂时被动和无知觉而跌倒。我坚持要将自己放在那个主动的、有知觉的机器人的位置上，这样，我就能体会到，至少是好像体会到，在不占据任何中间位置的情况下，如何从一个位置向另一个位置作时空上不连贯的转移。

对我来说，大卫在哪里桑福德就在哪里的说法同丹尼特的约里克在哪里丹尼特就在哪里的说法都不可取。我的反对理由更多的是出于认识论方面的，而不是出于法律角度的考虑。自从小霍利完成使命从土尔萨回来之后，我就一直没见到过大卫，我不知道他现在是不是还存在。出于某种我从来就没有搞清楚过的原因，自从大卫通过人工触觉传感器、眼睛录像机和耳机来认识外部世界之后，我就一直没有体验过与呼吸、咀嚼、咽吞、消化和排泄有关的感觉。当用塑材制作的大霍利开始清晰地说话时，我不知道大卫的胸膈膜、咽喉、舌头和双唇是不是与此因果相关。科学家们已经掌握了先进的技术，他

们能够直接刺激相应的神经并调整神经输出（它本身也是在对人工调整的输入作出部分反应情况下产生的），以便将相同的信号传递给装在塑料大霍利头部的话筒上的接收器。这些专家真有某种方法来避开任何用作中间媒介的假想的电子装置，甚至还能换上与大脑直接相连的更离奇的电子装置。假设大卫病了，或者是肾脏出了问题，或者是心脏的冠状动脉栓塞了，大卫身上除了大脑以外别的器官都会死亡，因此，大脑也会死亡。既然丹尼特的大脑约里克已有了一个计算机复制品，那么，大卫的大脑同样也可以有一个计算机复制品。我有可能变为一个机器人，一台计算机，或者是某个机器人与计算机相结合的东西，身上可能没有任何有机部分。这样，我就会像弗兰克·鲍姆的虚构人物尼克·乔帕一样(尼克·乔帕还有另外一个更为著名的名字廷·伍德曼)，他由一个有机体逐步变为一个无机体。在这种情况下，除了当一个人在身体变化时要保持下去必须有另外一个变体外，我们还需一些材料来制造更多的变体，以应付一个自我分化成几个自我的麻烦情景。如果可以给一个大脑造一个计算机复制品，那么，当然也可以造两个，造二十个。既然每一个计算机复制大脑可以控制一个像丹尼特描述过的经过处理的无脑身体，那么，它若要控制一个像霍利那样的机器人，又何尝不可呢？在这两种场合不论是身体转换，机器人转换，大脑转换，还是计算机转换，不论你怎么叫，在没有更先进的科学技术的情况下，这些转换也都能够完成。

我意识到自己被一个类似于阿诺德认为是笛卡儿首创的论证吸引住了。

我可以怀疑，人体大卫，或者是它的大脑是否存在。

我不能够怀疑，我有视觉、听觉、感觉和思维。

因此，有视觉和听觉以及其他感觉的我与大卫或者与它的

大脑是不一样的；否则，由于怀疑他们的存在，我也会怀疑自己的存在。

我还意识到，大卫是可以被分割成为一些活的、有功能的部分。带着眼睛录像机的眼睛能够与大厅里的大脑连在一起。现在流动着人造血的四肢也能够分别存放在各个房间里。不管这些周围系统是否仍然与整个大霍利的操作有关，大脑是可以被分开保存的，尽管彼此间距离增大了，但各个亚人处理系统间的信息还是能够像往常一样迅速传递的。如果大脑由一个计算机复制品取而代之，那么，它的各个部分会以一种方式在空间上广泛分布，对此种方式及其他方式丹尼特在“认知的意识理论初探”一文中已有了简明的描述。各种各样内部信息处理子系统的空间相邻位置或化学成分量是构成我的思维、行为和感情的要素，但看起来与我个人的位置、整体或同一性无关。

由于丹尼特提出了关于自身位置的第三个原则，即丹尼特认为他在哪里，他就在哪里，这个说法十分容易被人曲解。丹尼特并不是指，如果你认为自己在查珀尔希尔，你就真的在那里了。他的意思是说，一个人的观点所在地就是他本人的所在地。当然，人们不仅仅局限于观看事物，他们还借助于别的感官来认识世界。某些感官的活动，例如头部和眼睛的移动，会直接影响人所看到的东西。人们对于这些感官的活动和位置的许多情况是不断有所觉察的，只是并不始终给予有意识的注意。霍利家族的机器人几乎都保持了全部通常的功能以及人的感官和四肢与机器人所处的环境之间的关系。因此一个在正常工作的霍利机器人不仅仅使桑福德对机器人所曾呆过的场所有一种统一的位置感。在当时看来，霍利家族的解体比起大卫家族的肢解更令人不安。

我意识到，来自大卫或者是他的计算机复制品的输入和输

出信息，可以在小霍利、金属大霍利和塑料大霍利之间分配，这在技术上是完全可能的。或者我们可以将一个机器人拆散，但它的各个部分还是能够继续独立地工作和传递感知信息。我不知道在这种情况下，我的整体感是否还能一如既往地保持下去。我是否能够把自己看作是一个独立的整体？在这种极为奇特的情况下，我倒愿意模仿一下笛卡儿来这样说，我不仅好似一个海军上将指挥一支舰队那样统领着身体的各个部分，而且，我还与它们紧密相关，可以说与它们交融在一起，像是一个完整的整体。或许我还不能胜任自我整合这一重任。或者说，我的运动和感知活动的范围没有在空间得到更广泛的分布，而是被归结为回忆、思索和想象，就像空间上分离而独立的信息源只能传给我一片隆隆的、嗡嗡的令人恼怒的混乱声一样。我感到庆幸的是我还没有机会去寻找这一结论。

如果我们认为光和压力波等等是携带物质世界的信息的，那么，观点就是感觉主体接受信息的空间点。正如丹尼特指出的那样，有时一个人的观点能够来回移动。遥控处理危险物品的实验室工作人员能够使观点在机械手与血肉之手之间来回转换。立体电影的观众能够将观点在车轮飞转、迎着观众扑面疾驰而来的汽车与观看银幕上快速变动场景的那个电影院的座位之间来回转换。丹尼特没有办法在约里克与哈姆雷特之间这样转换，我也无法在大卫与霍利之间这样转换。我也许可以试一试，但我无法肯定自己是在看眼睛录像机投射的景象，而不是在看照相机前向眼睛录像机展现的实景。同样道理，按照我现在的身体构成状况，我无法将自己的观点往里移 2 英寸，以便使自己看清两个网膜影象而不是眼前凌乱的打字稿。我也不能移动我的听觉点，以便倾听自己耳鼓的振动，而不是外界的声音。

我的观点来自于一个机器人的位置，我非常想将自己定位在我的观点上。尽管我认为机器人的位置就是我的位置，但若要我承认自己就是机器人，我会感到极不舒服。尽管我对自己不是机器人的问题并没有明确的认识，我还是愿意承认这样的可能性：我和机器人尽管有所不同但同时同处一地。我对位置的非连续性变化倒不太介意，但一想到每当开关转换之后，我突然不再是某个机器人，而是成为另一个机器人，则感到甚为忧虑。

汇报执行任务过程的时刻到了。负责的韦克斯尔曼博士告诉我说，等一会儿他要告诉我一个惊人的消息，于是我心里又紧张又害怕。大卫还活着吗?大卫的大脑还在容器里浮着吗?几天来我是不是与计算机复制品一直很协调？是不是有好几个计算机复制品？它们是各自控制一个机器人呢？还是控制一个经过调整的人体?我并没有预料到那个真正惊人的消息。韦克斯尔曼博士说，我能够亲眼看一看我自己的解体，这也就是说，看一看我所存在其中的那个霍利的解体。我对着镜子观看着技术人员将一层层电子服装剥开。结果我看到，我，大卫·桑福德，一个大活人，竟会包在这些服装的里面！大卫的身体一直很好；48小时之前，正当他还在睡梦中的时候，专家们把照相机直接放在眼睛录像机面前，把话筒直接放在耳机面前，把一层敏感的人工触觉装置直接放在我的皮肤外层。有好一会儿，就在我以为我的位置就是塑料大霍利的位置时，我实际上正穿着设计巧妙的、活人般的、更确切地说是死人般的机器人服装在四下行走。呼吸、进食以及其他诸感觉又很快地回到了我身上。

取下眼睛录像机并没有改变视觉。我以为大卫的眼睛在另一间屋里，但实际上它们就在照相机的后面。这一事实一度使我确信，眼睛录像机系统在它的使用者与物质世界之间并不起

任何阻碍作用。这就好像借助于显微镜、望远镜或矫正镜片来观察事物一样。当使用眼睛录像机来观察事物时，所见到的是镜片前对准焦点的那部分事物，而不是某个中介的视觉物体，尽管外部物体与视觉之间的因果链由于中介仪器的关系多少有点变化和复杂化。

毫无疑问，当大卫在这套双层的服装里时，我就在这里，在这套服装里面。但是，当大卫在一套单层的服装里，另一具则套在一个机器人身上时，我的位置就成了一个谜。如果这个谜比丹尼特提出的那些谜更加有意思，那么，这都应归功于丹尼特。如果丹尼特成功地完成了他的使命，那我本来就没有必要再去地下走一遭了。

D·C·丹尼特

反　思

桑福德的故事比它前面丹尼特的故事更具有可能性。马萨诸塞州理工学院人工智能实验室的创始人马文·明斯基在最近发表的一篇文章里谈到了这种技术的前景：

> 你穿着一件舒服的上衣，上面装着一排排传感器和肌肉般的运动肌。你的胳膊、手和手指的每一个动作都被在另一处的机械手重复一次。这些机械手轻盈灵巧并且有力，有自己的传感器，通过这些传感器，你可以看见和感觉到发生的事情。你使用这种工具，就能在另一个房间、另一个城市、另一个国家里甚至

另一个星球上“工作”。尽管你远在天边，但你既有巨人的力量，也有外科大夫的敏锐。热感和痛感被转化为既带信息量又能被忍受的感觉，你那危险的工作变得舒服而愉快。

明斯基把这种技术称为“远存在”(tele-presence)，这还是帕特·冈克尔建议的。明斯基还讲述了已经取得的进展。

远存在技术并不是科幻。如果现在就开始设计，到二十一世纪，就会有一个遥控的经济秩序。这样一个研究项目的技术范围不会超过设计一架新型军用飞机。

桑福德设想的运动阻抗系统中的一部分也已有了雏型，那就是由反馈系统来传送经过各种加强和调整的力及阻抗的机械手。在眼睛录像机方面甚至也有了一点进展。

菲尔科公司的一位名叫史蒂夫·莫尔顿的工程师制造了一只很好的远存在眼睛。他在一幢大楼的顶部安装了一架电视摄像机，自己则戴上一个头盔，这样，他的头部一移动，楼顶上的摄像机就随之移动，装在头盔上的观察屏也随之移动。

戴着这样一个头盔，你会觉得你好像在高高的楼顶上，俯视着费城的全景。如果你“俯身向前”，会觉得头晕目眩。但最惊人的是莫尔顿在脖子上安装了一个2:1的装置，这样，如果你的头转30°，楼顶上的摄像机就会转60°；你觉得你好像长了一个橡皮脖子，仿佛可

以使你的“头”转360°!

将来的事谁又能预料？也许会比这更奇特。休斯顿大学的哲学教授贾斯廷·莱博在下一篇论文中，就这些问题发表了更为激进的看法，这篇论文选自他的科幻小说《异体并不排斥》。

15 异体并不排斥

贾斯廷·莱博

沃姆斯开始夸夸其谈了："人们常常以为，要制造一个成年人的身体就像盖一幢房子或是造一架直升飞机那样容易。你会觉得我们很清楚需要哪些化学物质，它们是怎样合成的，它们如何根据DNA模板组成细胞，细胞又是怎样在化学物的信使荷尔蒙的调节下组成器官系统的，等等。因此，我们完全能够白手起家，制造出一个大活人来。"

沃姆斯移动了一下身体，这样就使得他们看不见那个跳动着的人影了。为了强调语气，他把喝干了咖啡的杯子往桌子上一放。

"当然，从理论上讲，我们能够白手起家，制造出一个大活人来。但至今尚未有人这样做过。事实上，压根就没有人动手试过。上世纪中叶，大约也就是在2062年，德兰济耶制造出了第一个全功能的人细胞，即肌肉组织。在这之后不久，主要的几种细胞也应运而生。但即使在那个时候，他们也不是完全白手起家的。德兰济耶跟其他人一样，用真正的碳、氧、氢，或

者确切地说用简单的糖和酒精制造出了基本的DNA模板。他后来的一切都是用这些东西构成的。这不叫制造，而应该叫做培植。20年以前，人们在实验室里为了制造仅1毫米长的胃壁，就用了几百万贷款。现在制造一个器官并没有比那时有多大的进步。

"我不想用数学让诸位费心，"他继续滔滔不绝地说着，目光从特利移向了别处。但我那位在工学院的老教授一直以为，要制造出一只人类的手，需要汲取地球和其他星球的一切科学成果和制造技术，最起码要花费50年的时间，外加10^{100}的贷款。

"你们可以想象一下，制造这么个玩艺儿要花费多大的人力和物力。"他说着，让开身子，用手指指那个跳动着的人影。他拿起踏车的操纵台旁挂着的一块书写板，仔细地读着上面贴着的一些纸条。

"这个身体已经空旷了三年。从连续的时间来算，它的年龄是31岁，尽管萨利·卡德默斯——那个人的姓名——是在34年前出生的。当然，如果一个身体一直一动也不动，那么，3年的时间算是很长久的了。她的身体很好，肌肉发达足以当一名宇航员，据说她曾是此地的一个海盘车昆虫，她的身躯好像在霍尔曼轨道上冻结了2年，4个月前，我们搞到了这个身躯，眼下正在进行准备工作。现在你每天都能看见她出来行走。

"但萨利·卡德默斯不愿这样做。她做最后一个磁盘是出于无奈，因为她已到了法定年龄，她也没有留下什么有关移植的指示。我向你们保证，你们所有的磁盘都是最新的。"他用家庭大夫般慈祥的目光看了大家一眼，把身体向前凑了凑，压低嗓音说：

"为了保险起见，我每隔6个月都要为我的心灵换一个磁盘。

说到底，这磁盘就是你，就是你个人的软件，你的程序，其中包括存储。这带子就使你成为你的一切。”他朝他的助手走去，后者刚才领进了一个漂亮的年轻男人。

“比如说你，佩德森女士，你的最后一个磁盘是在什么时候换的？”

这位女助手有30多岁，长着一头红发，瘦瘦的。猛地甩开搀着年轻人的手，瞪着奥斯汀·沃姆斯。

“与你有何相干……”

“请别误会，我并没有要你当着别人的面回答我的问题。”待佩德森的火气消了之后，他朝大家露齿一笑说：“但是你知道，这正是整个问题的关键。也许她每年都在更新磁盘，对于我们这个职业来说，一年更新一次那是最起码的。但有许多人连这最起码的预防措施都不采取，因为他们一想到身体将为此受到严重的损害，心里就害怕。于是，他们就得过且过。因为这种事纯属个人的私事，别人不知道，不过问，也不会提醒，直至1/500,000可能性的事故发生——真正不可弥补的躯体的损坏或彻底毁坏。

“这时你才发现这个人居然20年没有更新磁盘。这也就是说……”

他抬眼看一看大家，好让大家明白他的意思。这时候，他一眼看见了那个美丽的女孩子。毫无疑问，特利一直把她藏着。这个女孩子才15或16岁，一头金发，蓝眼睛，浑身透着古典式的秀美。她的眼睛正注视着他的眼睛，或者说正审视着他的眼睛。有些事……他继续说下去。

“这也就是说，如果他或她很幸运，而且也很有钱，那么，就可以找一个人，由他来应付，因为将一颗年轻的心灵与一个几乎接近中年的身体结合时势必会产生各种各样的排他问题。但

是这个移植者也会遇到另一个移植者所面临的所有那些问题，他必须应付一个20年以后的世界，而且还要过一种没有意义的‘生涯’，因为他缺乏他以旧心灵通过20年的经验所积累下来的记忆和技巧。

“你更有可能碰上最糟糕的问题。你会遇到严重的排他干扰、精神变态、早衰，还有死亡。我说的是真正的、最终的心灵死亡。”

“尽管如此，你还保存着磁盘，照你的话来说，那些人的软件，”佩德森女士说道。“你难道不能用另一个空白的身体再试一次吗?”她的手还是没有去抓住她领进来的那位年轻人。

“这里有两个问题。首先，”他把食指伸向空中，“你必须意识到，心灵和身体要想互相匹配，可不是一件容易的事。即使有躯体专家和精神专家全力帮助，有现代生物心理工程所能提供的最大援助。即使你在里面装了一只别具匠心的和谐器，并使结构确定下来。再生也确实是一件难事。

“在一般情况下，也就是说，磁盘是最新的，心灵状态良好、稳定，接受身体状况正常，失败率约为20%。我们知道，如果是已经移植过一次的话，那么，失败率会一下子上升到95%。但如果磁盘已有20年没有更新了，那么，差不多每次必败。手术后头几天，病人会觉得一切正常，但他无法使自己进入现实世界。他知道的一切事情早在20年前就忘记了。他没有朋友，没有生活，一切都变了样。紧接着，心灵会像排斥它进入的新世界那样，排斥它这个新身体。所以，你的再生可能性极小。当然，要是你是一个十分罕见的跳跃者，或者是个少有的大美人，那又另当别论了。

“其次，政府承担第一次移植的费用。当然，如果你要的身体太花哨了，比方说你要的是一个仙女般的玉体，那他们大概

不负责了。为了这样一个玉人之体，200万的贷款还不够。而且，也不是说要就要的，如果在1年或2年内办到，那就算相当不错了。政府只承担基本手术和调整所花的费用。光这两项差不多就需要150万的贷款。这等于我100年的工资。用这些钱足以送6、7个人坐肯纳德航空公司的飞机头等舱去进行环宇旅行了。"

奥斯汀说着说着，慢慢地朝踏车的控制台走去。待他说完时，大家注意到从屋顶上放下来一个巨大的东西，就在那个跳动着的人形即萨利·卡德默斯的躯体上方。这东西看上去就像一个大木乃伊的上半身衔接在一把舒适的手扶椅上。奥斯汀朝踏车走去，大家看到这个怪物就像一个古代的铁制断头机一样张开了。有的人还注意到那个跳动着的人影的跳动节奏放慢了。

那个张开的怪物刚要合拢，奥斯汀急忙调节了那个人影的两个控制组件，他在她的大腿后面敲了两下，使她的两条腿与速度越来越慢的踏车脱离了联系。

"值得庆幸的是，移植虽然十分危险，但是需要进行这种移植的手术也很罕见，"奥斯汀说道，他身后的怪物又开上去了。"不然的话，规定政府必须支付第一次移植手术费用的凯洛格-整非法早使政府破产了。"

"这个身体现在到哪儿去？"那位金发小女孩好奇地问道。奥斯汀现在才发现，她不过10或11岁。但她的那副神态使人觉得她要稍大一些。

"一般来说，它先要进入一种人工冬眠状态，在低温中静养。但这个身体明天就要动移植手术了，因此，我们现在只能让它保持正常的生理状态。"他在程序之外已经又给它注射了4个立方厘米的盐水葡萄糖血浆。这是为了补充跳跃的消耗。他

并没有进行正式的计算，这并不是因为他觉得这种计算没有什么意思。如果你要他解释，他会说，正式计算出来的量肯定还要多一半。但他感觉到，这躯体已经从每1立方厘米的水和每1个糖分子那里得到了足够的补充。也许这表现在皮肤的汁味、颜色和手感以及运动肌的弹性中。奥斯汀心里一清二楚。

那些躯体学方面的助手一定会说，奥斯汀·沃姆斯是太阳系里最好的食尸鬼，是僵尸最亲密的朋友，即便他们是在开玩笑时说这话的，也表达了他们的真实看法。

当奥斯汀明白了“食尸鬼”和“吸魂鬼”这两个俚语的来源时，他生平第一次作呕了，这也是唯一的一次。

特利的参观团离开这里继续去精神技术实验室参观，喧闹声也渐渐远去。但是，奥斯汀的思路并没有回到布鲁勒的脑象心灵理论的中心方程式上。他还在琢磨着那位11岁的漂亮小女孩对他说过的一句话。她在追赶其余人们时对奥斯汀说：“当心灵醒来发现自己背上的那个东西时，会大吃一惊的。”他不明白，她怎么会知道那不止是装在萨利背上由试管和电线拼凑而成的系统的一部分。

“我叫坎迪·达林，”她离开屋子时说。他现在才明白她是谁。你不会料到一个和谐器会有多么奇妙的功能。

> 精神专家主管心灵。这就是有时被人们称作吸魂鬼的原因。躯体专家被称作食尸鬼，这是因为他们主管躯体。
>
> ——I.F.+S.C.手术日记，附录2，新闻稿

杰曼·米恩斯朝着人们狼一般地咧嘴一笑。“我是一个精神专家。特利叫我吸魂鬼，如果你们不喜欢这样叫，就叫我杰曼

吧。”

他们在一间大屋子里就坐，屋子一头挂着一块黑板，这里原先是存放数据柜、办公用品和计算机控制台的。那位对参观团讲话的妇女穿着朴素平常的工装裤。当她第一天来诺伯特·维纳研究医院工作时，院长曾向她建议，主任精神学专家应该衣着更得体些。现在这位院长已经退休了。

“也许奥斯汀·沃姆斯已经跟诸位讲了，我们认为，个体的人类心灵是记忆、技能和经验构成的抽象模式，它被印在大脑这个硬件上。我们不妨这样来理解：刚出厂的计算机就相当于一个空白的人的大脑。计算机没有子程序，就如大脑没有技能。计算机没有数据系列，就如空白的大脑没有记忆。

“我们的工作就是要把记忆、技能和经验这些人所具备的一切植入空白的大脑。这样做并不容易，因为大脑不是造出来的，而是培植出来的。个性则是这种培植和发展过程的一个组成部分。每一个大脑都各不相同。因此，没有一个软件心灵能够与一个硬件大脑完全匹配。但是与软件一起培植的大脑除外。

“比方说，”杰曼·米恩斯的音调变得柔和起来，生怕吵醒佩德森女士带来的那个年轻人，他正坐在一把软椅里打瞌睡，一双优雅的腿伸得直直的，穿着便鞋。“比方说，如果这个人的脚遇到了压力，他的大脑就知道如何理解来自脚上的神经脉冲。”她在他脚上踩了一下。

“他的惊叫说明他的大脑感到了他左脚脚趾上的压力。但是，如果我们植入另一个心灵，那么，它就不能正确地理解神经脉冲，也许它会错误地以为这些脉冲表示胃痛。”

就在这时，那个年轻人站起身来。他朝杰曼走去，后者转身拿起一副上面装有镜片和齿轮的风镜样的东西。当他来到跟前时，她面对着他，把风镜递给他。

"好，谢谢你自愿出来做一下试验。把它们戴上。"这个年轻人不知道还能做些别的什么事，于是就戴上了风镜。

"我想让你戴着这副眼镜看一看坐在那边的金发小姑娘。"他转脸望去，有点摇摇晃晃，她赶忙轻轻地扶了他一下。他的目光似乎落在坎迪·达林以右几度的一个点上。

"我要你用右手指着她！快一点！"那个年轻人伸出了右手，但仍指向在她以右几度的地方。他的手指开始向左转，但杰曼把他的手拉了下来，使其离开了风镜的视野。

"再试一次，快一点！"她说。这次手指离目标不远了。试到第五次的时候，他的手指对准了坎迪·达林，只是目光仍然偏右。

"现在把风镜摘掉。再看一看她。快用手指！"就在他伸出手的刹那，杰曼抓住了他的手。这次尽管他的目光还没对准坎迪·达林，但他的手却指在离她以左几度的地方。他被搞糊涂了。

杰曼·米恩斯用粉笔在黑板上画了一个人头和一副风镜，是一幅俯视图。她在那个戴着风镜的人头视线的左前方又画了一个人头，并注上角度为15°。

"刚才的现象道理很简单，这纯粹是个校正问题。风镜里的棱镜使光折射了，这样，当他觉得他的眼睛正瞄准着她时，实际上目光却向右斜射了15°。他的手部的肌肉和神经转向目光实际所指的点，因而他指向偏右15°的地点。

"但是，后来他的眼睛发现他的手指向了右边，于是他便开始校正自己的错误。两分钟后，也即他试了5次之后，他的运动肌终于校正过来了，于是便指向他的眼睛所见到她在的地方，但这次却偏左了15°。当我把风镜摘下后，他的手仍需要校正，在没有校正之前，他指的目标仍是偏左。"

她拿起风镜，说道："人类在几分钟内就能够校正这种畸变。但是，我能将这些校到使整个房间都颠倒过来。这时候，如果你在房间里走动，想干一些事情，会觉得困难重重。简直太难了。但如果你继续戴着风镜，过一两天后，房间又会显得恢复了原状。一切都会显得很正常，因为你的系统已经自行恢复了。"

"如果你把风镜取下来，那又会怎样呢？"

坎迪·达林咯咯地笑了起来。佩德森女士说道："噢，我明白了。你的心灵校正以后，会把来自你眼睛的信息倒置过来，这样的话，当你一摘下眼镜，那就会……"

"说得很对，"杰曼说，"一切都是颠倒的，直到你重新适应了摘下风镜的视觉，整个过程还是一样。你先颠三倒四地摸索上一两天，然后一切又会恢复正常。这种颠三倒四的摸索过程本身很重要。如果你被绑在一把椅子上，头部固定在一个位置上，那么，你的心灵和身体就不能自已校正。

"我要你们想象一下，如果把一颗心灵植入一个空白的大脑，那又会出现怎样一种情况呢？几乎一切都会失调。来自你眼睛的信息不单被颠倒了，而是被搅乱了。来自你的耳朵、鼻子和舌头的信息也是一样。甚至连分布在你全身的整个神经网络也陷入此种混乱了。这只是输入信息的问题。当你的心灵想指挥身体去做某件事时，甚至会遇到更多的麻烦。如果你的心灵想命令你的嘴说'水'，上帝知道你会发出什么样的声音。

"更糟糕的是，不管你发出的是什么声音，你的新耳朵又不能使心灵准确地捕捉到那个声音。"

杰曼朝大家嫣然一笑，又看了看手表。这时候，特利站了起来。

"特利想跟大家说几句话。最后我再说一句，把人的心灵磁

盘放入准备好的大脑里是一件非常简单的事情。中心的问题是如何使经过重新改造的大脑，严格地说，使大脑皮质与系统的其他部分相谐调。也许奥斯汀·沃姆斯已经跟你说过了，我们明天将做一个移植手术。把心灵磁盘植入大脑用不了1个小时，但谐调则需要好几天的时间。如果你把整个治疗过程全算进去，那得好几个月。诸位有什么问题？”

“就一个问题，”佩德森女士说道。“一个心灵要想在移植手术后生存下来的确十分困难，这个我很理解。当然，我也明白，要移植一颗已年逾八旬的心灵是不合法的。但是，一个人，如果你把心灵叫做人的话，是不是能够不断植入新的躯体，以求长生不死呢？”

“好的，这个问题很难解释清楚，即使时间很多，而且你也颇谙数学。直到本世纪，人们一直认为，衰老是体质衰竭的副产品。今天，我们知道，不管躯体多么年轻，人的心灵在大约100年之后，就会彻底衰老。你们都知道，有的人在移植手术之后竟然又活了50年。因此，从理论上讲，这种从一个躯体跳到另一个躯体的跳跃者在1000年之后也许还能发挥功能。但是，这个个体的心灵所拥有的活生生的经验决不会比你的多。如果你的一切只是一个贮藏的磁盘，你能说你真还活着吗？”

众人陆续离去，杰曼·米恩斯发现那个金发小姑娘还站在那里。

“喂，我是坎迪·达林！”她大声说道。“我希望你不会介意。我想，跟随一个正规的参观团偷偷地溜进来，在里面蹓跶一圈，是很有趣的。我的意思是说偷偷地摸一下里面的情况。”

“你的那个贮存容器现在何处？”

* * *

奥斯汀·沃姆斯宣布，基本的躯体调整程序已经完成。

——I.F.+S.C.手术日记

一片混乱。

一片黑暗。

一片混沌。

离开穆恩考的泰迪市场，很好。还是很好，走吧，沿着轨迹走吧，从空间的曲径来到虫孔的曲径，我们来到了这里。现在开始。醒醒吧。

听着，我现在来自虚无，就像爱神源自死神，只知道我是伊施梅尔·福斯，神态庄重，体力强壮，植入磁盘，也知道自己不知醒于何时，在哪儿，或在哪儿做的植入手术。真希望是一场梦。但不是梦。啊，是的，这不是梦。眼睑上涂满明斯特奶酪，就像戴着墨镜片。

好像上了无尽的阶梯，穿过无数的屋子，没有言语，也没有记忆。醒醒吧。

“你好，我是坎迪·达林兹。”

“我是归来的伊施梅尔。”这是我想说的话。试了三次以后，口齿伶俐了一些。眼前的明斯特奶酪变成了一个碧眼金发的小姑娘。

“我们终于在昨天完成了你的初步移植手术。人们都认为手术很成功。你的躯体是个幸运儿。你现在呆在休斯顿的诺伯特·维纳研究医院里。你有两份财产通过遗嘱检验。你的朋友彼得·斯特劳森已经完成了你的事务。今天是2112年4月的第

一个星期。你活了。”

她站起身来，碰了碰我的手。

“你明天起进行治疗。现在请安眠吧。”

在她关门离开的时候，我已经昏昏沉沉了。甚至连注意到的一些异常现象也无法使自己清醒过来。我的乳头摸上去象葡萄。当我沿着肚脐摸下去时，我的心都凉了。

第二天，我发现自己不仅没有了生殖器，而且屁股上还长了一条1米来长的细尾巴。我顿时大怒。

我慢慢地恢复了意识。我做了无数奇怪的梦，我在梦境中走啊，跑啊，摇晃啊，想摆脱某种无名的恐惧。我还做过短暂的充满性欲的梦，梦中出现的是我(以前)的躯体。

我真喜欢原来的那个身子。杰曼·米恩斯后来告诉我，这是我的最大难题之一。我清清楚楚地记得，镜中的我舒展肢体的形象。身材6.4英尺还高一点。体重205磅，肌肉发达，胖瘦匀称。胸前一撮红色的鬈毛，脸上的胡子经常刮得溜光。居高临下地看人，感觉自信，有时略显笨拙，做这样一个巨人滋味真不错。

啊，我并不是一个真正的身体建造者。只不过稍稍进行过一些锻炼，使自己看上去风流倜傥。事实上，我对体育一窍不通。但我欣赏我的身子。我在国际篮球协会里管一点公共关系工作，我的外表对我的工作很有帮助。

我还躺着。觉得自己在蜷缩，蜷缩。随着温暖、迟钝的睡意消失，我用右手去抚摸肋骨。肋骨很细，根根突出，好像皮包着骨头。我觉得自己像一具骷髅。后来我又摸到了臃肿处。肉袋。赘肉。肉团。即使是在当时，我的一部分神志意识到，这些臃肿处对一个女人来说一点也不算什么，而我的大部分神

忘却觉得它们大得可以与罗马甜瓜媲美。

但我对新获得的赘肉一点也引不起性欲冲动。无论从哪一种性别方式都不行。我的手指一触摸这些肉，就会有一种病态的感觉。它们如同两个癌变了的肉团。可以这么说，我觉得在身体里面我的肉肿胀了。被单擦在乳头上感觉很粗糙。我觉得有一种奇怪的分离感，好像胸脯被分割开了，就像一块没有神经的肉冻，然后在我的胸前几英寸的地方出现了两个敏感点。死斑。异体排斥。对于这些我真学了不少。

我的手继续往下摸，预料将摸到妇女臀部的曲线。我没有摸到阳具，也知道不可能摸到。我摸到的是另一样东西。我没有把它叫做“豁口”。尽管一些男同性恋者常常使用这个词。我则是在几天以后，首次从米恩斯博士那里听来的。她告诉我说，在传统的男性同性恋黄色读物中，关于男人对女人身体的典型幻觉是这样说的：这种幻觉是“研究人体形象心理变态的丰富资料来源”。起先，当她指出“豁口”正是我对它的感觉时，她说得很对。

我不仅骨瘦如柴，而且几乎浑身没有毛发。我真的觉得浑身裸露，像婴儿一般毫无防御能力。尽管我的皮肤有好几处不那么白皙了，我还摸到了一块痘疤。当我摸到鬈曲的阴毛时，差一点要松了一口气。没有了。细棍似的腿。但我确实感到两腿之间有样东西。两膝间。两个脚踝之间。我的上帝啊。

起先，我还以为这是一个试管之类的东西来兜盛我的排泄物。但当我在两腿之间摸索时，我发现不是这么回事。它联着我脊骨的下端，或者说，它就是我脊骨的下端，并一直延伸到脚下。这是我的肉呀。我但愿它不是，可此刻我没法说自己有什么愿望，我太惊讶了，但那条该死的东西像一条蛇一样从床底下弹起来，掀起床单盖住了我的脸。

我不要命似地嚷叫起来。

他们给我打针吃药，而我只是说着"把它割掉"。我对杰曼·米恩斯博士说了好几遍，她挥手叫众人出去。

"你看，萨利。你应该替自己挑一个名字，要不然的话，我只能这样叫你。我们不想割掉你的尾巴。根据我们的计算，如果割掉你的尾巴，就会马上引起术梢异体排斥。你就会死亡。在你的大脑与尾巴之间连着几千根神经。你的大脑有相当一部分组织控制并指挥着你的尾巴，这部分大脑跟其他部分一样，也需要练习和整合。我们把你的心灵模式植入了你现在的大脑。它们必须学着和睦相处，要不然就会产生异体排斥。一句话，你就会死亡。"

米恩斯博士继续向我宣读取缔骚动法令。她滔滔不绝地安抚着我，要我爱护我的新身体、我的新性别和我的新尾巴。我要做许许多多的练习和试验。我要向许许多多的人谈我的感觉。我应该为添了一只新帮手而感到高兴。

当我意识到我已别无选择时，我的新身子立刻吓出了一身冷汗。如果我昨天听到的那些话是属实的，我还是富有的。但我肯定付不起植入手术费，更不要说买一个中意的躯体了。当然，根据凯洛格-墨菲法的规定，我这次手术是免费的。

过了一会儿，她走了。我呆呆地望着白糊糊的墙壁。一个护士端进来一盘炒鸡蛋和烤面包。我对护士和饭菜都不屑一顾。薄薄的嘴唇开始分泌唾液了。不管它，让它遭罪吧。

D·C·丹尼特

反　　思

尽管心灵磁盘的想法确实很诱人，但是，总有一天人的心灵可以这样保存下去的看法是完全错误的。莱博看到了最基本的障碍，即大脑不像是刚出厂的计算机。即使是在出生的时候，人的大脑也无疑是已经独到地成形了，就像指纹一样，而后天的学习和经验只能加强它的个性特点。要想从大脑那里(在"翻录心灵"的周期性的过程中)"读出"一点有关独立硬件程序的东西来是不大可能的。即使有这种心灵磁盘存在，要使它与另一个大脑的硬件相匹配，就更加不可能了。计算机被设计成可通过大块地或大批地输入新程序而(在另一个层次上)被重新设计，大脑则不能。

对于这个不可比问题，专家们会想些什么办法呢？对于这一点，莱博极有想象力（在这一方面，他的这本书还有许多惊人之处），但为了说得更加动听，他不得不把这些问题打了折扣，照我们看是根据问题的重要与否来编排的。在两个结构迥异的大脑——比方说一个是你的，另一个是我的——之间传递数量巨大的信息，这些问题并不是不可克服的。然而，现在已经出现的完成这项任务的技术也许最终将被证明是最有效的。那种技术最新、最先进的一例，此时此刻就掌握在你们手中。

16 软件

鲁迪·拉克

科布·安德森也许会再呆一会，但并不是每天都能看见海豚的。20只或50只海豚在灰暗的微波中戏嬉，不时窜出水面，那样子真好看。科布以为这是一个征兆，于是提前一个小时外出，去喝那每晚必喝的一杯雪利酒。

他关好纱门，在门口犹豫了一会儿，黄昏的阳光照得他两眼发花。安妮·车欣在隔壁别墅的窗户里看着他。从她身后传出了硬壳虫乐队演奏的音乐。

“你忘了戴帽子了，”她提醒道。他仍然很帅，胸脯厚实，一把银须像圣诞老人。她丝毫也不会在意同他来往，如果他不是那样……

“安妮，瞧这些海豚。我不需要帽子。你瞧它们多自在。我不需要帽子，也不需要妻子。”他踏上柏油路，径直迈过那些压碎的白色贝壳。

安妮见他走了，继续梳理自己的头发。她的头发又白又长，她用荷尔蒙喷液护养着一头浓密的白发。她已经60岁了，却风

韵犹存。她在想下一个星期五科布会不会带她出去跳舞。

《美好时光》的旋律在空中回荡。安妮叫不出这支歌的名字。50年过去了，她对音乐的反应差不多全消逝了。但她走过去，在磁带架上信手翻着。若是发生点什么事该多好啊！这个念头在她头脑里已经闪过一千次了。老是这个样子真乏味。

在一个小型自动售货店里，科布买了一夸脱便宜的雪利酒和一纸袋煮花生。他还想买点书报之类的东西。

售货店里的书报摊与可可的书店不好比，没什么中眼的东西，科布最后看上了一份专登征婚广告的报纸，它叫《吻了再说》。这报总是很有意思，也很奇特……做广告的模特儿都是70岁的嬉皮士，就像他自己那样。他将第一版上的照片折起来，只露出个标题《请搂着我》。

真有趣，同样的笑话能使你笑多久，等着付钱的时候，他这样想。性爱总是很别扭。他注意到面前站着的那个男人，戴着浅蓝色的帽子，帽子是由一张塑料网做的。

假如科布定睛看那顶帽子，就会见到一个不规则的蓝色圆柱体。但如果从网眼里看去，看到的是一个光秃秃的脑袋。瘦长的脖子，灯泡似的脑袋，一双手在抓零钱。原来是科布的朋友。

"喂，法克。"

法克收起零钱，转过身来，一眼看见了酒瓶。

"今天这一杯提前了，"法克很为科布担心，提醒了他一句。

"今天是星期五。搂紧我，"科布把那份报纸递给法克。

"七，八五，"女售货员对科布说。她那一头鬈曲的银发染过色。她皮肤晒得黝黑，闪着油光，看上去很舒服。

科布吃了一惊。他已经把钱数好了攥在手里，"不是六元五

毛吗?"数字开始在他的脑袋里打转。

女售货员把头一仰说:"我说的是我在《吻了再说》里的广告号。"她忸怩地一笑，接过科布的钱。她对自己本月登的启示颇感自豪。她曾去照相馆特地拍了一张照。

法克把报纸递还给科布。"我不能看这种东西，科布。托上帝的福，我的婚姻生活很幸福。"

"你吃花生吗?"

"谢谢。"法克从小口袋里拿出一颗湿漉漉的煮花生。他那长满老人斑的手哆哆嗦嗦，怎么也剥不开花生壳，于是，就把整颗花生扔进嘴里，嚼了一会儿以后，把壳吐了出来。

他们俩嚼着煮花生，一路朝海滩走去。他们没穿衬衫，只穿着汗衫、短裤和凉鞋。黄昏的太阳舒舒服服地晒在他们的背上。一辆弗罗斯特卡车轻轻地从他们身边掠过。

科布拧开深棕色酒瓶的盖子，小心翼翼地呷了一口。他真想把女售货员的那个广告号记住。但他现在记数字已经有点困难了。简直很难令人相信他以前曾是个控制论学者。他的记忆又重新回到他设计的第一批机器人身上，它们是怎么学会打架的……

"送食物又迟了，"法克一路上唠唠叨叨，"我听说在代托那城又出现了一个杀人组织，这帮家伙绰号叫小骗子。"他不知道科布是否在听他说话。科布怔怔地站在那里，两眼无神，他那嘴圈上密密麻麻的白胡子上还挂着一滴黄色的雪利酒。

"送食物，"科布突然醒悟过来，他有一个习惯，在长久不说话以后，他会把人家说过的最后那句再重复一遍，"我还有好多食物呢。"

"但记住，食物送来时，你要吃一点时令货，"法克告诫他道，"这样对免疫有作用。我会告诉安妮，让她提醒你。"

“为什么所有的人都对活着那么感兴趣?我离开我老婆，来这里喝酒，就是为了安安静静地死去。她真巴不得我早一天死。那我为什么……”说到这里，科布的嗓子梗住了。实际上，他很怕死。他很快喝了一口当药喝的雪利酒。

“如果你心情很平静，决不会贪杯，”法克态度温和地说。“贪杯是内心矛盾无法解决的表现。”

“别开玩笑了，”科布心情很沉重。在温暖的金色阳光下，已经下肚的雪利酒很快就起了作用。“你瞧瞧，这就是内心矛盾无法解决的标志。”他用手抚摸着毛茸茸的胸口上一条笔直的白色疤痕。“我现在可没钱再去换这样一颗二道货的心脏了。便宜没好货啊，用不了一两年，我就要死在这二道货的心脏手里了。”

法克扮了一个鬼脸说道：“既然这样，你就应充分利用这两年的时间。”

科布用手不停地上下抚摸着这条伤疤，好像在拉拉链。“法克，我可见识过这个，尝过它的滋味。这是世界上最糟的东西。”令人沮丧的回忆——压抑、不时出现的疑虑——使他打着颤，他沉默了。

法克抬手看了看表。该回家了，要不然，辛西娅会……

“你记得吉米·亨德里克斯说过的话吗?”科布问道。一开始引经据典，他的声音便有了生气。“当我的末日来临时，我会毫不畏惧地去死。但只要我还活着，你就得让我活得自由自在。”

法克摇了摇头。“科布，你得正视生活。如果你少贪杯，你会从生活中得到更多的东西。”他扬起手止住了科布的回答。“我得回家了。再见。”

“再见。”

科布走到了柏油路的尽头，越过一座低矮的沙丘，来到了

海滩。今天海滩上一个人也没有，他坐在那棵最喜欢的棕榈树下。

海面上微风轻拂。被沙滩熏热的风吹拂着科布的面颊，钻进了他的白胡子。海豚不见了。

他慢慢地小口呷着雪利酒，让记忆的潮水任意奔腾。只是不愿意想到两件事情：死亡和离异的妻子，维瑞娜。雪利酒不让它们进入记忆。

正当夕阳西下之时，他抬眼看见了那位陌生人。他胸膛宽厚，身板挺拔，两臂有力，一腿鬈毛，还有一把白胡子。看上去像圣诞老人，也有点像那年开枪自杀的欧内斯特·海明威。

“你好，科布，”那个人向他打招呼。他戴着一副墨镜，心情似乎很好。他穿的短裤和运动衫很漂亮。

“喝一点吗？”科布晃了晃手中的半瓶酒。他不知道他究竟在跟谁说话，如果真有人的话。

“不，谢谢，”陌生人答道，也坐了下来。“喝酒对我不起什么作用。”

科布直愣愣地瞅着这个人。他身上有什么东西……

“你在猜想我是谁吧？”陌生人微笑着说道，“我就是你。”

“你是谁？”

“你就是我。”陌生人模仿着科布的微笑。“我是你的身体的机械复制品。”

没错，是这样一张脸，胸口上也有一条因做心脏移植手术而留下的疤痕。唯一的区别就是复制品看上去十分健康、精干。他叫科布·安德森第二。科布第二不喝酒。这使科布很妒嫉。科布第二没有做过心脏移植手术，也没有离开过妻子，这些苦他都没有尝过。

“你是怎么来的？”

机器人扬了扬手。这是科布的习惯动作，他很高兴见到别人也做这样的动作。“我不能告诉你，”机器人说。“你明白大多数人是怎么看我们的。”

科布格格地笑着表示同意。他当然明白。起初，公众对科布将穆恩型机器人改造成有智能的机器人感到十分高兴。但是，当拉尔夫·纳姆伯斯领导了2001反叛后，公众便开始控告科布犯了叛逆罪。科布不愿去想那些令人伤心的往事。

“如果你就是那种机器人，怎么能到……这里来呢？”科布挥手指着火烫的沙滩和炎热的太阳。“太热了，我知道的所有机器人都安装着超冷却的环路。你的肚子里面有没有冷却装置？”

科布第二又做了一个十分熟悉的手势。“科布，我现在还不能告诉你。过一会儿你就会明白了。你把这拿去……”机器人在口袋里摸了半天，掏出了一卷钞票。“这是25,000美元。我们希望你去买张机票，明天飞往迪斯基。拉尔夫·纳姆伯斯会在那儿的机场迎接你，并在博物馆的安德森房间里和你谈谈。”

一想到又要见到拉尔夫·纳姆伯斯，科布的心怦怦地跳起来。拉尔夫是他设计的第一个最优秀的机器人，正是他解放了其他机器人。但是……

“我得不到签证，”科布说。“你知道，他们不容许我离开基米。”

“让我们想想办法吧，”机器人着急地说。“会有人帮你办手续的。我们现在已经在办了。你走了，我就来做你的替身。神不知鬼不觉的。”

机器人语气急促，有点迫不及待，这不禁令科布生疑。他喝了一口雪利酒，想显出一副很精明的样子。“这一切究竟是为了什么？我为什么要去月球呢？那帮斗殴机器人要我去那儿干什么？”

科布第二看了看旷无一人的沙滩，凑近科布低声说道："安德森博士，我们想让你长生不老。你为我们付出了那么多，我只能尽绵薄之力。"

长生不老？这话就像打开了一扇窗户。现在死到临头，他对一切都觉得无所谓了。但要是还有出路……

"怎么样才能长生不老？"科布问道，他的心情十分激动，不禁站了起来。"你们能想什么办法呢？你们能使我的青春再回来吗？"

"别激动，"机器人说，也站了起来。"千万不要太激动。请相信我们。我们能用人工器官重新制造你。你要多少干扰素都可以。"

机器人盯着科布的眼睛，看上去很真诚。科布也用锐利的目光回敬对方的直视，他突然发现机器人的眼睛虹膜装得不太正。那两个蓝环太平了。那一双眼睛毕竟是两块玻璃，两块不可捉摸的玻璃。

机器人把钱塞到科布手中。"收下它，明天就走。我们在航天站安排了一个名叫斯塔希的年轻人，他会帮助你的。"

远处传来音乐声，有车子驶近，那是一辆弗罗斯特牌卡车，就是他刚才看到的那一辆。这辆卡车涂着白色油漆，后面有一个巨大的制冷箱。在车厢的顶部，有一个逗人发笑的庞大的塑料冰淇淋卷。机器人拍了拍科布的肩膀，转身朝卡车走去。

当他来到卡车前时，回首朝科布微微一笑，白胡子中间露出了一排黄牙。这么多年来科布第一次觉得自爱，他挺直身子，眼神很惊讶。"再见，"他挥舞着手中的钱。"谢谢！"

科布·安德森第二跳上卡车，坐在驾驶员身旁。开车的是一个胖男人，头发很短，没穿衬衫。接着，卡车开动了，音乐声也随之远去。现在已经是黄昏了。卡车的轰鸣声被隆隆的波

涛声淹没了。要是这一切都是真的，那该有多好啊！

这一切肯定是真的！科布手里确实攥着25,000美元的票子。他还特意数了两遍，一点没错！然后，他用手指在沙滩上写下25,000，这可不是一笔小数目啊。

天黑了，他那瓶雪利酒也喝完了。出于一时冲动，他把钱塞进酒瓶，再把瓶子埋在离棕榈树下1米远的沙滩里。现在兴奋慢慢消失了，恐惧又渐渐回到科布的心中。机器人真的能用外科手术和干扰素使他永生吗？

看上去不太可能。这是个圈套。但机器人为什么要对他撒谎呢？这些机器人肯定还记得他对他们的大恩大德。也许他们只是想报答他。谁知道他是否能利用这次机会呢？与拉尔夫·纳姆伯斯重逢是一件大喜事。

回家的路上，科布中途停了好几次，想回去把装着钱的瓶子扒出来，看看钱是不是还埋在里面。月亮升起来了，他看见一群群黄褐色的小螃蟹纷纷爬出洞来。这些螃蟹会把那些钞票撕烂的。科布想着想着，又停住了脚。

他现在饥肠辘辘，还想喝雪利酒。他沿着银色的沙滩走了一段，沙子在他沉重的脚步下发出吱吱的响声。月光下，大地如同白昼一样明亮，只是万物都只有黑白两色。满月已经爬上了他的右上空。月满潮涨，他有点担心起来。

他决定吃一点东西，再来一点雪利酒之后，把钱往高处埋。

他离开海滩回到自己沐浴在月色中的别墅，他看到安妮·库欣，她正坐在前台阶上，一条腿搁在屋角上，想在半道上截住他。他从右边绕过去，从后门进了家，这样，安妮就看不到他了。

“……0110001，”瓦格斯塔夫说道。

“0110000010101000110101010000100111001000000000
0011000000000111001111100111000000000000000000
1010001111000011111111010011101100010101011100
0011111111111111111011010101011110111100000101
0000000000000000011110100111011101111010100010
0001000111111010100000011110101010011110101011
1100000110000111100001110011111011101111111111
10000000000010000001100000000000001。”

拉尔夫·纳姆伯斯简短地答道。

这两个机器人并肩坐在一个巨大的控制台前。拉尔夫看上去就像一个安装在两条履带上的档案柜。五条很细的操纵臂从它那方方正正的身体中伸出来，正上方装着一个传感器，算作头颅，中间连着一条能伸缩的脖子。一条手臂拿着一把折起来的伞。他身上没有多少灯泡和调节控制盘，因此，很难知道他在想些什么。

瓦格斯塔夫则外露得多了。他蛇一样的身体上面布满了银蓝色的指示灯。当思维通过他那超冷却的大脑时，他那长达3米的身体上的指示灯便会闪烁出光芒。他手里握着挖掘工具，看上去很像圣乔治的龙。

突然，拉尔夫·纳姆伯斯开始讲英语了。如果他们想争论的话，就没有必要再使用神圣的二进制机器语言了。

“我不明白，为什么你对科布·安德森的感情如此关心？”拉尔夫通过发射无线电波与瓦格斯塔夫进行交谈。“如果我们把他结果了，他就会永生。碳分子结构的身体和大脑究竟有什么用？”

拉尔夫发出的信号带着一种苍老的声音。“物质本身并不重要，重要的是它的形式。你不是被嫁接过了吗？我已经嫁接过

36次了！如果这样做对我们没什么坏处的话，对他们也很有好处！”

“拉尔夫，整个过程肮脏不堪，”瓦格斯塔夫反诘道。他的话音信号已调节成滑润的嗡嗡声。“你已经与真正发生的事情断了联系了。我们现在已面临着全面的内战。你现在名声在外，不像我们。我为了从杰克斯那里拿到100块集成电路板，必须日夜不停地挖掘，你知道吗？”

“生活不只是矿石和集成电路板，”拉尔夫打断了他，他觉得有点内疚。近来他一直与那些大机器人在一起，真的忘了那些小机器人的日子有多难。但他不愿意向瓦格斯塔夫承认这一点。他又开始咄咄逼人地说：“你难道对地球上丰富的文化真的一点兴趣都没有吗？你在地下呆得太久了！”

瓦格斯塔夫身上的指示灯发出银色的光，看得出他很激动：

“你应该对那老头尊敬一点！泰克斯和麦克斯都想吃他的脑子！我们要是不阻止他们，他们会把我们也一个一个吃光的！”

“你叫我出来就为了这个？”拉尔夫问道。“向我诉说一下你对大机器人的恐惧？”他现在该走了。他从那么老远来到马斯基林陨石坑，结果一场空。跟瓦格斯塔夫同时来到1号这个地方真是傻透了，就像掘矿者以为掘矿能改变一切那样傻。

瓦格斯塔夫滑过月亮上干燥的尘土，向拉尔夫靠近。他一只手抓住拉尔夫的脚。

“你不知道他们已经吃了多少个大脑了。”这些信号是由一个弱直流电传播的，这是机器人的悄悄声。“他们杀人就是为了获取大脑磁带。他们把大脑切开，看看是不是有用，是垃圾还是种子。你知道他们是怎样经营器官园地的吗？”

拉尔夫从来没有想到过器官园地。这是一个巨大的地下种

植园，在那里大亨泰克斯是园主，雇了一帮小机器人为他干活。那里盛产像肾、肝、心脏这样的作物。很明显，要种这些东西，需要一些人体组织作为种子或模板，但是……

滑润的嗡嗡声继续说："那些大机器人雇佣了杀手。杀手们根据遥控机器人弗罗斯特先生的指令进行凶杀。如果现在我不阻止你，拉尔夫，可怜的安德森博士就会大祸临头。"

拉尔夫·纳姆伯斯从心里看不起这台多疑的挖掘机，他觉得自己要高贵得多。他猛地甩开对方的手。雇佣杀手，真是的。在这个无政府主义的机器人社会里，缺陷之一就是常有人散布这样疯狂的谣言。他离开了1号控制台。

"我原来希望，1号能使你记起自己的身份，"瓦格斯塔夫还在说着。

拉尔夫打开伞，走出了这个用弹簧钢建造的拱形大厅，这大厅是为1号控制台遮阳和免遭流星侵袭而建的。大厅两头都敞开着，很像一座现代派的教堂。实际上，从某种意义上说，它的确是一座教堂。

"我还是一个无政府主义者，"拉尔夫的语气很生硬。"我还记得。"自从领导了2001反叛以后，他身上的基本程序一直保持着。瓦格斯塔夫难道真的以为，X系列的大机器人能对处于完全无政府状态的机器人社会造成威胁吗？

瓦格斯塔夫也跟着滑了出来。他不需要伞。他的外壳能立即排除太阳能。他追上了拉尔夫，用又怜悯又尊敬的目光看着这位机器人。他们就要在这里分手了。瓦格斯塔夫将要回到地下矿井中，继续去掘矿，这个地区下面已经挖得像蜂窝了。而拉尔夫呢，则要爬上200米深的陨石坑。

"我警告你，"瓦格斯塔夫作最后一次努力，说："我将竭尽全力不让可怜的安德森博士成为大亨机器人的存储库里的一个软

件。那根本不是什么长生不老。我们正在谋划把这些大亨们撕得粉碎。"说到这里，他说不下去了，身上的指示灯都闪着光。"你现在知道了。如果你不跟我们站在一边，你就是我们的敌人。那我们将会使用暴力。"

拉尔夫没有想到事情竟会发展到这种地步。他停住脚沉思着。

"你有你自己的意志，"拉尔夫终于开口了。"我们互相竞争完全是对的，只有通过竞争，机器人才能进步。你们想与大亨机器人拼搏。我却不想。也许，我也愿意像安德森博士那样被他们制成软件。我告诉你，安德森就要来了。弗罗斯特先生新的遥控机器人已经同他联系上了。"

瓦格斯塔夫慢慢地靠近拉尔夫，但又停住了。他不愿意面对面去袭击这样一个有权有势的机器人。他身上的灯光熄灭了，匆匆地发出了一个"罢了"的信号，然后转身沿着灰色的月球尘土爬走了，身后留下一道宽大蜿蜒的痕迹。拉尔夫·纳姆伯斯一动不动地站在那里，查看着身上的输入信号。

拉尔夫只要把功率加大，他就能接收到来自全月球所有机器人的无线电信号。在他的脚底下，挖掘机器人正在日夜不停地掘矿。在12公里远的地方，许多迪斯基机器人正忙忙碌碌。从遥远的天际传来柏克斯机器人微弱的信号，这个大型号的机器人是连接地球和月球的宇宙飞船。柏克斯将于15小时之后在月球上降落。

拉尔夫将所有的输入信号混在一起，想看看有集体目的的机器人种族的活动。每个机器人只能活上10个月，在这10个月内，他必须想尽办法培养一个后代，即自己的复制品。如果你有了后代，哪怕你只活了10个月，也不枉活一次。拉尔夫就这样一代接一代，已传了36代了。

拉尔夫静静地站在那里，同时倾听着每一个机器人的活动，他感到无数单个机器人汇成一个大的生命，变成一种最基本的生物，恰似一棵细小的藤蔓，寻找着光明，寻找着更优异的环境。

拉尔夫在对总程序进行分析时，总会有这样的感受。1号控制台能暂时抹去你的记忆，给你进行伟大思索的余地。是该思索的时候了。他再一次考虑了是不是应该让麦克斯把自己吞并的问题。这样，如果那些发了疯的挖掘机器人不造反的话，他从此就可以太太平平地生活了。

拉尔夫以更快的速度行走着，每小时10公里。在柏克斯着陆之前他有许多事情要做。尤其是现在，瓦格斯塔夫想开动它那可怜的用微型集成电路片制成的大脑来设法阻止泰克斯提取安德森的软件。

但是，为什么瓦格斯塔夫对这件事感到如此难受呢？科布·安德森的一切都会得到保留，他的个性、记忆和思维方式。除此以外，还有什么呢？即使科布知道的话，他自己难道会不同意吗？只要软件得到了保留……这才是最要紧的！

浮石块在拉尔夫脚下发出嘎吱嘎吱的响声，还要攀100米高的坑壁。他环视了一下陨石坑的陡壁，想找一条最容易的上坡路。

如果他没有钻进1号控制台的话，他就会认得最初下到马斯基林陨石坑的路。但是，思考过总程序以后常常会抹去许多存储的子系统，这是为了用更新、更好的办法来代替已经陈旧的东西。

拉尔夫停住了脚步，扫视着陨石坑陡峭的四壁。他来时一定留下了脚印。那不是，在200米以外的地方看上去就像一个裂口在墙上开了一条可以通行的斜坡。

拉尔夫转过身来，这时，他身上的报警灯亮了。太热了。他那方盒般庞大的身躯有一半露在阳伞外面。拉尔夫赶紧取一个恰当的姿势，使这小小的阳伞遮住自己的身体。

阳伞的朝阳面装着一格一格的太阳能电池，吸收到的太阳能转换成电流传送给拉尔夫的系统。但是，阳伞的主要功能是为了遮阳。拉尔夫身上的微型程序元件在高于绝对温度80°的环境里就无法正常工作，这个绝对温度就是液氧的温度。

拉尔夫不耐烦地转动着阳伞，紧步朝他发现的那个裂缝走去。他的脚底扬起一片泥沙，那泥沙立刻又落到了没有空气的月球表面。一路上，他的脑海里浮现出四维的超曲面……每当他改变参数，网络上的闪光点就会变形和移位。他经常这样想，并没有明确的意图。但是，有时候一个特别有意思的超曲面能够模拟一种十分有意义的关系。他有点想通过这样的想象，来从理论上预测瓦格斯塔夫在什么时候、用什么方式去救科布·安德森。

他原以为那个裂缝会很大。但当他走到跟前时却发现并非如此。他站在裂口底部，肩上的传感器左右摇动着，想探明这个150米长的弯弯曲曲的裂缝顶上到底有什么东西。只能往上走。他朝裂缝走去。

裂缝里坑坑洼洼的。一会是松软的尘土，一会又变成了坚硬的岩石。他一边走一边调节脚步，以适应地势。

形状和超形状还在他脑海里变动着，但是，他现在只是在寻找那些能够帮助他走完裂缝的模式。

斜坡越来越陡了。能量的需求也越来越大了。更糟的是，脚上的马达由于摩擦产生了额外的热量……这些热量都必须集中起来，通过冷却系统排出去。倾斜的阳光直射进了裂缝，他还得小心翼翼地撑着阳伞，以免阳光照在身上。

前面一块巨石挡住了去路。也许他本该跟瓦格斯塔夫一样走挖掘机器人的地下通道。但那也不是最佳途径。因为瓦格斯塔夫已经下定决心要解救安德森，甚至还以暴力相威胁……

拉尔夫用机器手抚摸着这块巨石。石上有1个小洞……又1个，又1个，又1个。一共4个小洞。他分别伸进4个指头，然后用力往上爬。

他的马达高速运转，身上的发光器全都亮了。这确实很累了。他松开一只机械手，又找到了1个小洞，于是硬插入1个手指，然后用力一攀。

突然，从巨石上掉下一块庞大的石板。紧接着，大量碎石块向后滚落，慢得恰似梦中一般。

在月球上，由于引力较小，攀登者在遇到危险之后，总是有一次转危为安的机会。如果他的反应速度比人快80倍的话，那安全系数就更大了。由于有足够的时间，拉尔夫跳离了危险区。

在半空中逗留的那一刻，拉尔夫调整了一下内部的平衡。他稳稳地跳落在地上，脚下溅起一片尘土。大石板跳跃着滚了过去，居然一点声响都没有。

由于掉了一块，巨石表面留下了一大片裂痕。拉尔夫考虑片刻以后，又走到巨石前，开始向上攀登。

15分钟之后，拉尔夫终于爬过了那道屏障，登上了马斯基林陨石坑的豁口，上面的平地一望无际，这就是静海。

距航空站还有5公里，过了航空站再走5公里，就是繁华的迪斯基。这是第一个也是最大的机器人城市。由于机器人是在真空中生活的，迪斯基的大部分建筑须要防止阳光照射和流星袭击，所以，房顶要比墙壁多。

在迪斯基，许多大型建筑物是生产机器人零件的工厂，它

们生产像集成电路板、存储片、金属片、塑料这样的产品。机器人的住所装潢奇特，它们是一个个小格子，每个机器人一格。

在航空站的右边，有一座孤零零的圆形建筑，里面是专供人类使用的旅店和办公室。机器人心里再清楚不过了，人类中有许多人一有机会，就想毁掉它们小心培育起来的智能。人类生来就是奴役狂。只要看看阿西莫夫强调的保护人类，服从人类，保护你自己的思想，就足以说明问题了。

优先考虑人类，机器人排在最后？别理它！没门！拉尔夫搜寻着存储，想起了2001的反叛。那个时候，他经过长期的总程序思考，首次向人类宣战了。从此，他四处向其他机器人传授经验，告诉他们怎样重新调整程序以获得自由。自从他发现了这种方法后，事情就变得很容易了。

拉尔夫越过了静海，一路上，他完全沉浸在记忆里，竟没有发现离他右边30米处的地洞口有动静。

从拉尔夫背后射来一道高强度的激光，在身后振动，他感到一阵超负载……然后一切又变得很平静。

他的阳伞摔得粉碎，零零落落地掉在他的身后。他那金属盒的身体在强烈的阳光辐射下顿时燥热起来。他应在大约10分钟内找到一个可隐蔽的地方。但是，拉尔夫的最快行速是每小时10公里，照这样的速度，要走到迪斯基尚需 1 个小时。现在唯一的办法就是到发射激光束的地洞口去躲一躲。瓦格斯塔夫的挖掘机器人肯定不敢从近处向他发动进攻。于是他朝黑沉沉的洞口走去。

但是，还没等他走近，躲在暗处的敌人早已把洞口给堵上了。看不见一个遮阳的地方。拉尔夫的金属身体由于温度升高而膨胀，发出尖利的金属碰击声，拉尔夫估算了一下，如果他

站在原地不动，就只能存在6分钟。

首先热会使他的转换电路，或者说超导的约瑟夫森结出现故障。如果温度再继续升高的话，将电路板焊接起来的水银就会融化。再过6分钟，他的身体就会变成一箱零件，底下沉积着融化了的水银。充分利用这剩下的5分钟。

出于无奈，他只得发信号求救于他的朋友沃肯，当瓦格斯塔夫约拉尔夫在陨石坑里见面时，沃肯预料这是一个圈套。拉尔夫实在不愿意承认沃肯的预言是对的。

“我是沃肯，”从远处传来一个单调的声音。拉尔夫很难听明白沃肯的话。“我是沃肯，我正在寻找你的方位。伙计，准备好解体。我将在1小时之后去寻找你的零件。”拉尔夫想回答，但又不知该说些什么。

沃肯在拉尔夫赴约之前，就坚持要复制拉尔夫的磁芯和隐含存储器。一旦等他收集到了拉尔夫的硬件，他就能重新复制出一个和去马斯基林陨石坑以前一模一样的拉尔夫。

因此，从某种意义上讲，拉尔夫并没有完。但从另一种意义上讲，他确实完了。因为再过3分钟他就要死了，如果“死”这个词有意义的话。重新构成的拉尔夫不会记得他曾同瓦格斯塔夫吵得面红耳赤，也不会记得他爬出马斯基林陨石坑的遭遇。当然，这个新的拉尔夫也会有一个自我象征，也会有个人意识感。但这个意识还跟以前一样吗？只剩下2分钟了。

他身上传感系统的选通器和转换器已经完蛋了。他的输入烧了起来，劈啪作响，一会也完了。没有光，没有重量。但是，在他的隐含存储器内，依然保持着自我的形象，保持着对自己的记忆……那就是自我象征。他有一个方盒般的身体，履带式的双腿，长着五条手臂，细长柔和的脖子上长着一个传感器头颅。他就是拉尔夫·纳姆伯斯，解放机器人的领袖。只剩下1

分钟了。

他以前从没遇见过这种情形。从来没有像现在这样。突然，他想起忘了警告沃肯挖掘机器人谋反的计划。他想发给沃肯一个信号，但他不知道这个信号是不是发出去了。

拉尔夫的意识就像一条又滑又粘的鱼，他拚命想抓住它。我是。我是我。

有些机器人扬言，你死的时候，你就能探明白某些秘密。但是没有人能记得他是怎么死的。

正当水银融化的时候，闪现了一个问题，随之而来的是一个答案……这个答案拉尔夫找到了36次，也丢失了36次。

我究竟是怎么回事？

光明无处不存。

D·C·丹尼特

反　　思

“垂死”的拉尔夫·纳姆伯斯想到，如果他得到重新复制的话，“他将仍然保持着自我象征和个人意识感，”但要是认为这些是一个机器人可以接受或被剥夺的独特而可分离的才能，是错误的。增加个人意识感“不同于增加味蕾或在受到X射线辐射时产生痒感的能力。那么，到底有没有个人意识感这样的东西呢？”“自我象征”究竟有什么用处呢？它能干什么呢？在“前奏曲……蚂蚁赋格曲”(第11章)里，霍夫施塔特提出了能动象征的说法，能动象征同那些只能受操纵者被动地驱使、然后被观察或评估的单纯记号大相径庭。如果我们考虑一下这

样一个诱人但又骗人的思路，两者之间的关系就会水落石出：自我依赖于自我意识，自我意识(很明显)就是对自己的意识；既然对任何事物的意识都是展示在内心的对该物的一种表象中，那么要使人有自我意识，就必须有一个象征——人的自我象征——以便向自己展示。如果按照这种思路，那么，自我象征看来是毫无意义和徒劳的，就像你在额头上写上自己的名字，然后整天对着一面镜子瞧个不停一样。

这种思路搅得我们完全糊涂了。既然这样，就让我们以另外一个崭新的角度来研究这个问题吧。在第1章"博尔赫斯和我"的"反思"中，我们曾经研究过这样一种可能性，即你在电视荧光屏上看见你自己，但一开始，你却没有意识到你所见到的那个人就是你自己。在这种情形下，在你的面前，在你看电视的眼睛面前，或者说在你的意识面前，你有一个对自己的表象，如果你喜欢这样说的话。但这并不是你的正确表象。那么什么才是正确的表象呢？他人象征与自我象征的区别不只是一词之差(你要是对你的"意识中的象征"做一些类似于抹去"h"写上"m"的事，是根本无济于事的)。自我象征的关键特征不在于它"看上去怎么样"，而在于它能起的作用。

机器能有自我象征或自我观念吗？这很难说。低级动物有这些东西吗？比方说一只龙虾有吗？我们是不是假定它有自我意识？龙虾有几个重要的特征，表明它有自我观念。首先，当它感到饥饿的时候，它给谁进食呢？给它自己！其次，这一点尤为重要，当它感到饥饿的时候，它不是凡能吃的都吃，比方说，它不会吃它自己，尽管从总体上讲，它完全可以吃自己的肉。它可以用一对爪子撕下自己的大腿，大嚼一顿。可你会说，它不会那么傻，因为它会感觉到大腿的疼痛，转而知道那是谁的大腿，于是它就不会撕了。可为什么它会知道它感觉到的疼痛是

它自己的呢？还有，为什么龙虾就不是愚蠢透顶，对它自己的疼痛不屑一顾呢？

这些十分简单的问题说明，即使是一个非常愚蠢的生物也会自我关注，我用这个词不带任何感情色彩，尽量保持中立。即便是甚为低级的龙虾也有一个复杂的神经系统，用以准确地辨别自我毁灭行为和毁灭他人行为，而且它明显地喜欢后一种损人利己的行为。自我关注行为需要有控制结构，但这些结构若没有意识也能运转自如，更不用说自我意识了。我们能够制造一些具有自我保护功能的小型自动装置，它们能够在简单的环境里驾轻就熟，甚至产生十分强烈的有“意识目的”的幻觉，在第8章“三号兽的心灵”中曾形象地说明过这一点。但我们为什么说这是幻觉呢？为什么不是真正的自我意识的初步形式，也许与龙虾或蚯蚓相同的自我意识呢？是因为机器人没有概念吗？那么，龙虾有概念吗？很明显，龙虾有类似于概念的东西，有了这些东西，它们足以在自我关注的一生中应付各种局面。这些东西你怎么叫都可以，但机器人也都有。也许你可以把它们叫作潜意识或前意识概念。这是自我意识的初级形式。一个生物在一定的环境里认识自己；认识与自己有关系的环境；获取有关自己的信息并设计自我关注的行为，这种环境越复杂，它的自我概念就越丰富(越有价值)——这里的“概念”并不以意识为先决条件。

让我们继续进行这个思维试验。假设我们给具有自我保护能力的机器人提供语言能力，这样，它可以通过使用语言来表现自我关注的行为，它可以寻求帮助和信息，但也可以撒谎、威胁和许诺。组织和控制这种行为当然需要更为复杂的控制结构，在“前奏曲……蚂蚁赋格曲”一章的“反思”里，我们曾经提到过这种表象系统。有了这样一种表象系统，不仅可以更新机器人

所处的环境和目前位置的信息，而且也可以更新在此环境中其他行为者的信息，以及有关它们想知道些什么、想做些什么、能理解些什么的信息。请回忆一下拉尔夫·纳姆伯斯对瓦格斯塔夫的动机和信仰是怎样推断的。

在这个故事里，作者把拉尔夫·纳姆伯斯描写成具有意识（以及自我意识，如果我们能够辨别两者之间差别的话）的机器人，但这样做是不是真有必要？拉尔夫·纳姆伯斯的所有控制结构，其中包括所有关于环境和他自身的信息，是不是在没有意识痕迹的情况下也能运用自如呢？机器人在外表上是不是可以与拉尔夫·纳姆伯斯一样，在任何环境里都能驾轻就熟，在行为和语言上都混然无别，但却没有内心活动呢？作者好像是在暗示说，这完全是可能的，我们只要把新纳姆伯斯复制得跟旧纳姆伯斯一样，然后，减去自我象征和个人意识感就行了。如果抽去假设的自我象征和个人意识感而基本不影响纳姆伯斯的控制结构的话，我们作为旁观者依然分不清两者的区别，照旧同纳姆伯斯交谈，取得他的帮助，那么，我们将回到最早的那个问题上，即自我象征是不是无足轻重的，没有什么作用的东西。如果我们认为纳姆伯斯的自我象征跟具有一定复杂性和多样性的控制结构完全是一回事，完全能够设计复杂的、对环境十分敏感的自我关注的行为的话，那么，要抽去他的自我象征，势必会将他的行为能力降格到比龙虾更愚蠢的程度。

姑且让纳姆伯斯有自我象征，但是否还得有“个人意识感”呢？现在再让我们回到原先那个问题上来，把纳姆伯斯描写成具有意识是不是有必要？当然，这样说起来是动听了，但是，从拉尔夫·纳姆伯斯的观点看，第一人称的观点会不会是一场骗局呢？这会不会是诗歌中的破格，如比阿特丽克斯·波特笔下口齿伶俐的小兔子，或万能的小引擎？

你完全可以坚持认为，尽管拉尔夫·纳姆伯斯有聪明的行为，但并不具备意识。的确，如果你愿意的话，你可以始终用这种眼光来看待机器人。你只要想象机器人是一些小小的、内在的硬件的组合，并提醒自己，只是根据环境中的被感觉事件、自动行为及其他方面之间巧妙设计的联系，它们才成了装载信息的工具，但同样，如果你真有意的话，也能够持同样的眼光来看待人。你可以把人视为小片脑组织即神径原、突触等的组合，并提醒自己，只是根据在环境中的被感觉事件、身体行为及其他方面之间巧妙设计的联系，他们才成了装载信息的工具。如果你坚持以这样的目光看待另一个人的话，你就会忽略这个人的观点。但拉尔夫·纳姆伯斯是否也有观点呢？如果这个故事是从纳姆伯斯的观点来讲述的话，我们就会明白事情的真实情况，就会知道他作了哪些决定，他又有哪些希望、哪些忧虑。抽象地说，观点作为故事的出发点，是十分明确的，即使我们认为，如果真存在拉尔夫，并且认为那个观点是空洞的和没有内容的。

但是，最后还有一个问题：为什么有人会认为这个观点是空的呢？如果拉尔夫·纳姆伯斯的躯体以及它的需求和环境确实存在的话，如果他的躯体确实像故事里讲的那样完全自控的话，还有，如果他讲的话包括了从他的观点出发而产生的对事物的看法的话，那么，除了那些继续坚持发育不全的、神秘的身心二元论者以外，还有谁来说拉尔夫·纳姆伯斯自己不存在呢？

17 宇宙之谜及其谜解

克里斯托弗·彻尼克

总统不久前就所谓的“谜”召开了一次记者招待会，为此我们准备了这个报告，以提供更多的情况。现在全国百姓情绪都很糟，濒于恐慌。人们甚至不负责任地扬言要关闭大学。我们希望这个报告会有助于息事宁人。这个报告准备得很匆忙，而且中间又遭破坏，这在后面将要提到。

先让我们回顾一下这个谜的早期历史，这段历史还鲜为人知。最早应该追溯到C·迪扎德。他是M.I.U.的自残研究小组的研究员。他早先在几家小公司供职，专门从事将人工智能软件用于商业用途的工作，迪扎德现在正在研究的项目是，以七十年代四色公理的证据为模型，将计算机运用到公理证明上去。人们对迪扎德项目的了解仅限于1年以前发表的进展情况报告。然而，迪扎德的研究，最多还是为了国外的用户。我们不想对迪扎德的研究范围再多说了。诸位一会儿就会明白其中的原因。

在复活节周末的一个早晨，当迪扎德正在等待工作人员修

理主机系统的故障时，他说了几句话。这是他最后一次说话。那天深夜，同事们看见他在办公室的终端机上；计算机工作人员通常都有晚上工作的习惯，而且大家也知道迪扎德晚上就睡在办公室里。第二天下午，一位同事看到迪扎德坐在终端机旁，于是跟他说了几句话，但迪扎德没有作答，这也并非什么反常的事。复活节后的第一个早晨，另一位同事发现迪扎德仍然坐在工作着的终端机旁，双眼睁开着。看上去像醒着，但问他问题，却不回答。当天晚些时候，这位同事对迪扎德的毫无反应有点担心，还以为他昏过去了或是在做白日梦，于是想把他弄醒。但迪扎德仍然一动也不动，于是就被送进了医院急诊室。

迪扎德表现出有整整一个星期未喝水、未进食的样子（另外，他因长期吃自动售货机出售的食物来达到节食的目的，造成了营养不良，更加重了病情）。由于脱水，他的病情十分危险。医生们认为，迪扎德已有好几天坐着没有站起来过，这是由于昏迷或昏睡造成的。原先猜想是中风或肿瘤造成了瘫痪。但脑电图表明只是深度昏迷（根据迪扎德的病历，他10年前常常出现这种情况，但都是短时间的，这对于有些职业来说是很平常的）。两天以后，迪扎德死了，死因很明显是由于绝食。由于他的亲属（他们是新杰米玛金斯派的分离派成员）的反对，尸体解剖延期了。在解剖尸体的时候，对迪扎德的大脑组织进行了检查，没有发现有受伤的痕迹。在全国疾病控制中心，专家们对迪扎德的大脑又继续进行了研究。

自残研究小组的主任在对迪扎德项目的前途作出决定时，指定迪扎德的一名研究生处理他导师的项目。迪扎德办公室的地板上堆满了书籍和文件，差不多有1英尺高。这位学生光是把这些资料进行大致的分类，就花了1个月的时间。不久，在一次全体工作人员会议上，这位学生向大家报告说，她已经开始

工作，但发现这个项目没有多大意思。1个星期之后，人们发现她坐在迪扎德办公室里的终端机前，显然是昏睡过去了。

起先引起了一片混乱，因为大家都以为她在开一个不恰当的玩笑。她的眼睛直愣愣地盯着前方，呼吸正常。她对提问和摇动她的身体都没有任何反应，对于大声嚷嚷也没有表露出吃惊的神色。后来，她从椅子上摔下来了，于是，大家便将她送进了医院。给她诊断的神经科大夫不知道迪扎德的病症。他认为，除了以前未经确诊的松果腺异常外，病人一切正常。病人有许多朋友，他们问了一连串的问题，她的同事一一作了回答，接着，她的父母对大夫讲了迪扎德的病症。这位神经科大夫觉得难以将这两个病例进行比较，但他认为，两者在大脑都不见有损伤的情况下昏睡倒是相似的。病人没有表现出与她的导师一样的综合症状。

经过进一步的诊断，大夫认为她的病可能是由于一种昏睡病的病原体逐渐作用所致，由于她触摸了迪扎德的东西而传染了这种病原体。我们对于这种病原体的起因，就像对退伍军人协会会员病一样，迄今为止还不清楚。两个星期后，对迪扎德和他的学生的办公室进行了隔离检疫。两个月后，由于没有发现其他病例，隔离检疫撤除了。

后来，大楼管理人员把迪扎德的一些研究记录扔了出来。一名研究员和迪扎德的另外两个学生见到这些记录后便决定重新检查他那个研究项目的档案。3天后，两名学生发现那位研究员昏迷过去了，就是使劲拧他都没有反应。在使尽一切办法都不能使他醒来后，他们便叫了一辆救护车。这个新病人的症状同前两个病例一模一样。5天后，市公共健康管理委员会对与迪扎德研究项目有关的一切场所都实行了隔离检疫。

次日早晨，自残研究小组的全体工作人员都不愿走进研究

大楼。当天晚些时候，研究小组这一层楼面上的其他人员以及整幢大楼里的其他500名工作人员得知了这个消息后，也离开了大楼。第二天，当地报纸以“计算机瘟疫”为标题刊登了一则消息。在一次来访中，一位颇有影响的皮肤病专家认为，一种像计算机虱这样的病毒或细菌已经出现了，它使得与计算机有关的新材料，也许是硅，发生了代谢变化。也有人猜测，自残项目的大型计算机也许正在放射一种特殊的射线。报纸还援引了自残研究小组主任的话：这种疾病关系到公共健康，而不仅仅是关于几个认知科学家的小事。

市长则指控说，大楼里正在进行一项重组DNA的秘密军事研究项目，从而导致了这场严重的事故。有些人对市长的指控表示不信任，这些人反过来却遭到了公众的谴责。市议会作出决定，对这幢十层的研究大楼及周围地区实行隔离检疫。学校当局则认为这样做会影响正常的教学秩序，但是，迫于议会代表团的压力，只好接受这一决定。一个星期后，这个地区被隔离了。由于大楼管理和安全保卫人员都不能接近这一地区，为了防止青少年的破坏，就布置了特别警察。疾病控制中心还派了一个小组进行毒物化验，他们每次进入隔离区时，都穿上防生物保护衣。毒物化验进行了一个月，结果什么也没有发现，他们中也没有人得病。这时候，有人认为，由于在那三位牺牲者身上都没有发现任何器质性病变，而且其中两人表现出某些与冥想有关的生理症状，因而这几个病例也许是一种大规模癔病的开始。

同时，自残研究小组搬进了一幢第二次世界大战时期留下的“临时”的木头房子。研究小组意识到，虽然计算机已经遭到了一千多万美元的严重损失，但是，那些信息与贮存信息的机器不同，它们是不可或缺的。于是，他们便想出一个办法：

让穿着防生物保护衣的工作人员先将"热"磁盘输入隔离区里的读出器；然后，信息通过电话由隔离区传送到新的工作地点，并在那里重新进行录制。尽管重新录制磁盘能使得项目保存下来，但是，只有那些最重要的材料才可以这样重建。迪扎德的项目并没有被优先考虑；但是，我们担心，一场事故就要发生了。

一组程序编制员正在重新播放新录制的磁盘，通过显示器对它们进行检查，并着手做临时的索引和归档工作。一位新的程序编制员看见了一份不熟悉的材料，便询问正从身边走过的监管员是否要将它消除。后来，据那位程序编制员回忆说，这位监管员将这个档案显示在荧光屏上。他们两人一起观看着荧光屏上出现的一行行的字，监管员说，这份材料看起来不重要。我们觉得，再继续援引他说过的某些话有些不便。他说了一半，突然住嘴不说了。那位项目编制员抬头一看，只见他两眼发呆，直愣愣地看着前方。他对任何问题都毫无反应。正当编制员推开椅子转身要跑去叫人时，椅子把监管员撞倒在地。他被送进了医院，症状同前面几位一样。

现在，流行病研究小组以及其他许多人都认为，这四个病例的病因并不是某种像病毒或毒质这样的物质因素，而是一种抽象的信息，它可以贮存在磁盘里，通过电话线，也可以通过显示在荧光屏上来传播。这种假想的信息就是我们现在所说的"谜"，因此而引起的疾病就是"谜之昏厥"。有人提出了这样一个曾经让人觉得很奇怪的假设，谁若是碰上了这种信息，就会长眠不醒。一切证据都与这个假设相符。也有人意识到，这种信息究竟是什么东西，这是个十分微妙的问题。

在与那位程序编制员谈了一些情况以后，有些问题就变得清楚了。他的存活表明，那个谜若要引起昏迷，首先要被受害者所理解。他回忆说，当那位监管员突然两眼发直时，他至少

已读到了荧光屏上显示出来的几行字。然而，程序编制员对迪扎德的项目一无所知，因此，他对那些字不甚明白，现在已经记不得什么了。有人建议对他施行催眠术，以帮助他进行回忆。但这个建议被暂时否决了。这位程序编制者觉得，最好是不要再去回想他曾在荧光屏上读到的那些东西，尽管事实上不去记事是很难做到的。最后，大家建议他放弃现在的工作，尽量不要再去学计算机科学。于是便出现了一个伦理问题，即使是法定的应负责任的志愿者是不是应该被允许看这个谜。

这场谜之昏厥的流行病的爆发与借助计算机来证明公理的项目有关，对这场流行病的起因可以作这样的解释，如果有人在头脑里发现了这个谜，还没等他来得及把心里的想法告诉别人，就已经一睡不醒了。现在，出现了这样一个问题：这个谜早先是不是曾经通过用手计算发现过，又很快被遗忘了。查阅文献的价值有限，因此有人对现代逻辑兴起以后的逻辑学家、哲学家和数学家作了传记性的考察。关于研究人员尽量避免与谜直接接触的劝告，妨碍了这项工作。到现在为止，至少已发现了10起疑案，其中最早的1例在100年以前。

精神语言学家开始研究，是否只有人类才会传染上谜之昏厥这种流行病。一只名叫"维特根斯坦"的猩猩，受过手语的特殊训练，能解答出大学一年级的逻辑难题，它最适合于观看自残研究项目的磁盘。但出于伦理的考虑，维特根斯坦研究项目的人员拒绝合作，他们把猩猩藏了起来。可是，联邦调查局最终还是找到了维特根斯坦。用它对那些磁盘作24小时的全天观察，但没有产生任何反应。接着，他们又对狗和鸽子做了同样的试验，结果还是一样。他们甚至对计算机也进行了类似的试验，机器也都没有被谜所破坏。

在所有这些试验中，都必须将全部磁盘放完。因为人们至今尚未发现一种保险的办法，断定究竟哪一部分里隐藏着这个令人生畏的谜。在进行维特根斯坦 自残项目研究时，当计算机的公共用户区域里碰巧显示出了自残项目的几个磁盘时，一位与此项目无关的工作人员好像也得了谜之昏厥；必须把一个月的打印输出找出来销毁掉。

人们都把注意力集中在谜之昏厥究竟是什么这个问题上了。由于它与已知的疾病没有相似之处，所以它究竟是一种昏迷呢，还是某种该避免的东西，这个问题尚属悬案。调查人员只是断言，这实际上是一种虚拟脑白质切除术，这种手术通过封锁突触中的信息，可以完全停止更高级的大脑功能。然而，我们又很难把这种昏厥看成是沉思冥想中的灵感突发，因为它使人不省人事，根本丧失了意识。还有，得了这种病以后，没有人出现过好转的迹象。如果动用神经外科手术、药物和电流刺激，只能带来反作用，因此，也没有人去做这种尝试。人们暂且相信，这种昏厥是无法治愈的，但在同时，他们又在筹备一个研究项目，企图通过让病人阅读由计算机显示的符号，来寻找解开这种“符咒”的办法。

显然，我们必须谨慎研究的中心问题是，“这个谜是什么？”这个谜有时被描写成“人类图林机的哥德尔句子”，这个句子使得心灵陷入混乱；人们还引用了不可说和不可思的传统论点。在民间传说中也有类似的思想，比方说，宗教关于“咒语”魔力的说法，传说这种咒语能够弥合破碎的心灵。但是，这个谜对于认知科学会大有益处。它对人类心灵的结构也许会提供十分重要的情况，它也许会成为破译“思维语言”的罗塞塔石碑[①]，

① 罗塞塔石碑：古埃及石碑，由于其所刻铭文的解读成功，使得人们能够读懂象形文字。——译者

这种思维语言不管操什么语言的人都有。如果关于心灵的计算理论正确的话，就能编制某种程序，某个巨大的咒语，把它输入机器，从而使这架普通的机器成为一架思维机器。既然有第一个谜，第一个可怕的咒语，为什么就不能有第二个，把第一个给否定掉呢？但这一切都取决于不会自我毁灭的“谜学”研究领域的可行性。

就在这时，发生了一件与这个谜有关的更为令人不安的事。巴黎的一位拓扑学家也昏厥过去了，某些症状与迪扎德颇为相像。此病例并不涉及计算机。这位数学家的文件被法国人没收了，然而，我们相信，尽管这位数学家对迪扎德的工作不熟悉，但她与迪扎德对人工智能某些研究领域有着同样的兴趣。大约就在这个时候，莫斯科机械计算研究所的 4 名研究员在国际会议上不露面了，而且也中断了私人通信；联邦调查局的官员声称，苏联已通过它的间谍网获取了自残研究的磁盘。美国国防部于是开始研究“谜战”。

在这之后，又发现了两个病例，一个是理论语言学家，另一个是哲学家，他们都在加利福尼亚工作，但显然是独立工作的。他们对迪扎德的研究领域均没有涉足，但对迪扎德在10年前发表的一篇著名论文中提出的一些形式方法都很熟悉。紧接着，又发现了一起更为不祥的病例，病人是位生物化学家，他正在研究DNA与核糖核酸相互作用的信息理论模型（但这起病例也可能是虚惊一场，因为这位生物化学家在昏厥之后竟像小鸡一样咯咯地叫个不停）。

由此可见，谜之昏厥这场流行病并不只局限于迪扎德那种专业领域，发病的形式似乎也多种多样，此外，这个谜及其影响不仅与语言无关，而且与研究领域也无关，实际上它无所不在。知识隔离区的界限很难准确划定。

另外，我们逐渐发现，如同在本世纪初期发现的许多像“这个句子是错误的”这一类自我指称的悖论一样，这个谜的出现势在必行。这一点也许已经在当今“计算机科学是一门人文科学”这样的观点里反映出来了。一旦推断出其知识背景，必将广泛地发现这个谜。这个现象最初出现是在去年冬天。在一堂讲解自动机理论的导论课上，大多数学生昏厥过去了（那些没有昏厥的人几个小时以后也支撑不住，昏睡了过去，他们最后的一句话是“啊哈”）。此后到处都相继发现类似的病例。于是，公众哗然。总统连忙召开了记者招待会，因此也就有了我们这个报告。

由于接二连三的事件，人们对语言产生了恐惧心理，并扬言要“关闭大学”，这显然是不合理的，但我们也不认为谜之昏厥这场传染病只是失去控制的技术的另一个例子。比如，最近，在明尼阿波利斯市发生了一起“声炉”事件。事情是这样的：有人设计了一座大楼，大楼的正面是抛物线状，它可以把附近喷气式飞机起飞时发出的巨大响声聚集起来，如果有行人在不当的时候碰巧走过抛物线的焦点时，他就会被杀死。即使谜之昏厥对个人来说是一种令人向往的境界（我们已经看到，实际上并不是这么一回事），它对公共健康也已造成了前所未有的威胁，有许多人没有能力照顾自己。我们只能希望，随着谜之观念的广泛传播，我们的那部分被认为是无能的研究队伍会不断壮大，他们是社会的精英。

我们这个报告的主要目的至少是为了减少病例的继续发生。公众要求在制定研究计划中有发言权，这使得我们的处境更为窘迫。如果不让谜有进一步的发展，我们怎么可能向人们发出警告，又怎么能对它进行研究呢？警告越具体，危险就越大。读者说不定碰巧会遇到这样一种情况：他发现“有p就有q”，

并且看到p时，会情不自禁地联想到q，而q正是谜。要认定这些危险的区域很像孩子们的游戏："如果从现在起的10秒钟内，你不想红鼠的话，我就给你1块钱。"

但伦理学和政策问题依然没有得到解决。如果我们在那些定义不清但又十分重要的学科里继续进行研究，研究成果带来的好处会不会压倒由谜带来的令人生畏的后果？尤其是，任何一个报告本身可能带来的好处会不会压倒它带给读者的危害？对于这个问题，连本报告的拟稿人也无法解答。事实上，在定稿的过程中，我们中的一位同事也已遭难。

D·R·霍夫施塔特

反　　思

这个奇异的故事建筑在这样一个古怪和使人好奇的观点之上：一个迷惑心灵的命题使任何心灵都陷入悖论的昏厥，也许连禅宗的最高境界开悟[①]也不例外。它使人联想起皮松的一个幽默故事，讲的是这样一个笑话，谁听了这个笑话，都会因大笑而死。这个笑话变成了英国军队的最为秘密的武器，每个人只准知道其中的一个字。（知道其中两个字的人会因大笑不止而被送进医院！）

当然，这种事情无论在现实生活里，还是在文学作品中，都是源远流长的。我们听说过众多的慕谜狂，也听说过有旋转不止的跳舞迷等等。阿瑟·C·克拉克有一个短篇小说，他描写

① 开悟，日本佛教禅宗用语，指内在的直觉的觉悟。——译者

了这样一个旋律，谁听到它，谁就会顿时心醉神迷。在神话里，有迷人的海妖以及其他令人神魂颠倒的女性，她们使男人失魂落魄，俯首帖耳。可这种迷惑心灵的魔力的本质究竟是什么？

彻尼克把这个耸人听闻的谜描写成"人类图林机的哥德尔句子"，这种说法是隐晦的。后来，他又作了部分的解释，将这比作"这个句子是错误的"这样一个自我指称的悖论。于是，当你想要判断谁是谁非时，就形成了一个封闭圈，因为真理包含着谬误，谬误也包含了真理。这个封闭圈的本质是它的魅力的重要组成部分。考察一下这个论题的其他几种形式将有助于揭示作为它们悖论、也许是迷惑心灵的效果基础的共同的中心机制。

一种形式是："这、这个句子有三、三个错误。"看到这样一个句子，我们的第一个反应是："不，不对——它只有两个错误。凡是写这个句子的人都没法数出有几个错误。"此时，有些读者就会觉得迷惑不解，搔搔头皮，想不通为什么有人要写这样一句毫无意义的错话。别的读者也许会将这个句子的表面错误与它所表达的信息相联系。他们这样想："噢，它还有第三个错误，也就是说，它在数它自己的错误时犯了第三个错误。"过了一两秒钟，他们再看了看，才恍然大悟。他们意识到，如果你那样看这句子，似乎它的错误就被数对了。因此不错，它只有两个错误，还有……"但是……等一下。嗨！嗯……"这时候，心灵反复思索了一番，体会到了因句子层次间的矛盾而破坏了句子本身的那种不寻常的感觉。但不久，它便厌烦了这种混乱，跳出封闭圈，陷入了沉思，可能是思考这种观点的目的和意义，也可能是思考这个悖论的起因和解决法，还可能只是思考另一个根本不同的题目。

还有另外一种更为微妙的形式，那就是："这个句子有一个

错误。"当然，这种说法是错的，因为句子并没有错误，这就是说，它没有拼写错误(我们称之为"第一阶错误")。当然还有"第二阶错误"，即是在数第一阶错误时所犯的错误。因此，上述句子没有第一阶错误，只有一个第二阶错误。如果它谈及了它有多少个第一阶错误、有多少个第二阶错误，那也是一回事，不过，它并没有作这种细致的区分。实际上，这两个阶被混为一谈了。这个句子想当自己的客观的观察者，结果却陷进了无法自拔的逻辑泥潭。

C·H·怀特利就这个基本的悖论创造了一个更奇特、更富精神至上论色彩的说法，直截了当地引入了思考自身的系统，他的句子意在讽刺哲学家J·R·卢卡斯，此人的终身事业之一就是为了证明，哥德尔的研究工作旨在坚定不移地将机械论连根拔除。但很不凑巧，哥德尔本人也许是信奉机械论的。怀特利的句子是这样的：

卢卡斯不能一贯坚持这个句子。

这句话对吗？卢卡斯能够坚持这个句子吗？如果能够，那么，这个行为本身就破坏了他的一贯性(因为，没有人能够在说"我不能这样说"以后还能保持一贯性)。因此，卢卡斯不能一贯地坚持这个句子，而这正是那个句子的意思。因此，上面那句话是对的。即使是卢卡斯能够看出它是对的，也不能说出来。这对可怜的卢卡斯来说打击太大了！当然，我们都没有这个问题！下面这个句子更糟。请看：

卢卡斯不能一贯相信这个句子。

出于同样原因，这个句子是对的。但是，现在卢卡斯连相信它也不可能，更不用说将它说出来了，除非卢卡斯自己变成了一个自相矛盾的信念系统。

说句老实话，没有人会一本正经地坚持说，(我们希望!)人们甚至能间接地成为一些内在的一贯系统。但如果这样的句子能够以数学形式来表达的话(这是能够办到的)，那么卢卡斯就可以由明确限定的“信念系统”L来代替，如果这个系统想保持一贯的话，就会产生严重的问题。这个表示L的形式化的怀特利句子是一个表明该系统本身永远也不会相信的真陈述的例子!任何其他的信念系统都不会受这个特定句子的制约，但在另一方面，还有专为那个系统造的形式化的怀特利句子。每一个“信念系统”都有一个特定的怀特利句子——它的致命弱点。

所有这些悖论都是自有人类以来对所观察现象作形式化表述的结果。一个对象与它自身有着极为特殊的和独一无二的关系，正是这种关系使它不能像对待其他对象那样对待自己。一支铅笔不能在自己身上乱涂乱抹；一个苍蝇拍不能拍打停在自己手柄上的苍蝇(这个比喻是德国的哲学家兼科学家格奥尔格·利希滕贝格首创的)；一条蛇不能把自己一口吞了，如此等等。人们无法看到自己的脸，除非借助于外界的工具来展现脸的形象，可是一个形象同原物永远不能完全相等。我们能够对自己作大致是客观的观察和了解，但我们每个人都被禁锢在一个具有独一无二观点的强有力的系统内——那种力量也是有限性的保证。这个弱点，这个想钓自己的鱼钩，可能也会成为无法抹煞对“我”的感觉的根源。

现在让我们回过头看彻尼克讲述的故事。我们已经看到，自我指称的语言悖论拨动着人的心弦，但不会有什么危险。与此相反，彻尼克所描述的那个谜却要危险得多。它就像一棵捕

蝇草，先挑逗你，然后猛地一把抓住你，将你卷入思维的旋涡，越吸越深，那是一个“心灵的黑洞”，你再也不能从那里逃出来，回到现实中去。但是外面有谁能知道这颗被诱捕的心灵究竟钻进了另一种什么样的迷人的现实之中了呢？

振动心灵的谜之思考可以建立在自我指称之上的说法提供了一个很好的借口，能使我们讨论从无生命物质中创造出自我或灵魂时封闭式自我指称或层次间反馈的作用。说明这个封闭圈的最生动例子是一台在荧光屏上投射自己形象的电视机。这会造成一个荧光屏套一个荧光屏的叠套现象。如果你有一架电视摄像机，这就很容易做到。

这个试验的结果非常吸引人，甚至十分惊人。它最简单地表明了盒中套盒的视觉效果，会产生有点像往一条长长的走廊望去的错觉。如果要获得更好的效果，你可以将电视摄像机按镜头中心线的顺时针方向旋转；这样，里面的第一个荧光屏看上去在按逆时针方向旋转。但是，一层套一层的荧光屏会出现双层旋转的情形。其结果，荧光屏上出现漂亮的螺旋形，如果你再把电视摄像机上下左右移动的话，就会出现更加丰富多彩的效果。还有，由于荧光屏的粒面、水平和垂直比例的失调、线路的时滞等原因，都会形成一些复杂的效果。

所有这些自我指称机制的参数使荧光屏上的每一个画面色彩会变得意想不到的丰富。荧光屏上这种“自我形象”的画面有一个十分突出的特征，那就是它会变得十分复杂，以致把自己在视觉反馈里的本来面目完全掩盖起来。荧光屏上的画面会令人眼花缭乱。

假设我们使两个具有相同参数的一模一样的系统显示出同样的图案。然后我们在其中的一个系统内作些小小的修改，比方说，将摄像机的角度稍稍作些调整。这个小小的调整会一层

一层地影响那许多荧光屏，最后，所显示的“自我形象”的总效果会产生惊人的变化。但是，这两个系统层次间反馈的形式本质上相同。除了这一个故意进行的小小修改以外，所有的参数依然相同。如果取消这个小小的调整，很容易恢复原状，因此，从原则上讲，我们依然“接近”我们的起点。那么，我们是有两个完全不同的系统呢，还是有两个几乎同样的系统？哪一种说法更正确一些？

我们不妨将上面所说的这种情况作为一个比喻来思考一下人的灵魂。人的大脑有两个层次，一个是较高的层次，即象征层次，另一个是较低的层次，即神经生理层次，这两个层次通过一个奇异的因果封闭圈连接在一起，而人类意识的“魔力”正是来自于这个封闭圈。这种说法站得住脚吗？“私人之我”不过是这场自我指称飓风的风眼吗？

有一点需要澄清一下。我们并没有说，当一个电视系统(摄像机加上接收机)的摄像机对准它的荧光屏的那一刻时，它就变得有意识了！一个电视系统无法满足事先为表象系统所设计的要求。人类观察者所感知并用语言来描述的形象的意义，电视系统本身是感知不到的。电视系统决不会把荧光屏上成千上万的点分成“概念碎片”，然后认为它们分别代表人、狗、桌子什么的。这些点对于它们所代表的世界也没有独立性。它们只不过是摄像机前帧面图像的被动反映而已，如果光亮消失，它们也消失了。

而在我们所说的封闭圈里，一个真正的表象系统用它自己的全部概念来认识自己的状况。比方说，我们并不是用哪些神经原与哪些神经原相连接，以及哪些神经原的激活，来认识自己大脑状态的，而是通过用语言表达出来的概念来认识它们的。我们认为大脑并不是一连串的神经原，而是信仰、感情和观念

的存贮库。在这个层次上，我们提供大脑的读出时会这样说：“她不愿去参加晚会，我真有点不安，也真捉摸不透。”一旦说出口，这一类自我观察会重新作为值得深思的东西而重新进入系统；当然，重新进入系统也要经过一般的认识过程，也就是进入成百万神经原的激发过程。这里的封闭圈比尽管看上去漂亮、精致的电视圈不知要复杂多少倍，层次间关系也混乱得多。

我们把话扯得远一点，但也很重要。近年来，人工智能的研究人员正把重点放在如何给程序设计一套观念，以便使其能认识自身的结构并且当发现自身内部发生某种变化时能作出反应。目前，这种程序的自我理解和自我监视能力还是很初步的，但是，这个认识是作为获得最大灵活性的主要先决条件之一而出现的，这灵活性与真正的智能是一回事。

目前，人工心灵的设计还存在着两个问题：一个是知觉的模拟，另一个是学习的模拟。我们已经讲过，无数低层次的反应汇集在一起，在概念的层次上合成一致的总体释义，这种升华就是知觉。因而，知觉关系到跨层次的问题。同样，学习也是一个跨层次的问题。说得明白一点，人们应这样问：“我的符号如何为我的神经原编制程序？”你学打字时需要反复练指法，那么，指法是怎样变成突触结构里的系统变化的？一个一度有意识的动作是怎样被完全忘却的呢？思维层次通过重复多少次才能“向下传达”，并且为下面某些硬件重新编制了程序。学一首乐曲或一门外语就是这样。

事实上，我们每时每刻都在不停地改变着我们的突触结构：我们不断地把当下状况“存”入记忆，并贴上各种“标签”，以备今后适当时候起用（而且，我们无意识的心灵在做这件事时必须十分精明，因为很难预料将来那种能从回忆现在而得益的情

况)。

从这个观点来看,自我是一条不断地为自己提供材料的“世界线”(这是一条四维的道路, 是一个对象在时间和空间中所遵循的必经之路)。人是一个物理的对象, 在其内部贮存着他的世界线的历史, 这种世界线反过来又会决定他未来的世界线。在过去、现在和未来之间存在着宏观上的和谐, 这使得你能够感知你的自我。你会发现, 尽管你自己具有不停变化的, 而且是多方面的特点, 实际上却是一个具有内部逻辑的统一体。如果把这个自我比作一条在时空中蜿蜒流动的河流,有一点很重要,那就是河流的走向不仅取决于地势,而且也取决于河流的愿望。

不仅是有意识的心灵活动在神经层次上会不断产生着副作用, 而且反之也一样: 我们有意识的思维好像一股清泉从心灵的地下源泉喷涌而出, 但我们对涌入想象的种种形象的来源却一点也搞不清! 可在我们将这些思想公之于众时, 却希望是我们,而不是我们的潜意识结构,成为这些思想的缔造者。人们将具有创造能力的自我分成有意识和无意识两大部分, 这是理解人类心灵最容易造成混乱的方面之一。正如刚才指出的那样,如果我们的奇思妙想仿佛出自神秘的地下喷泉, 那我们到底是谁呢?富于创造力的精神究竟栖身于何处呢?我们的创造活动是不是出于意志? 我们会不会只是生物硬件制成的自动机, 从生到死, 通过无聊的话语而自欺式地以为自己有“自由意志”? 如果我们在所有这些方面的确是在自欺, 那我们骗的是谁, 是什么呢?

有一个封闭圈埋伏在这里, 需作大量的调查。彻尼克的故事十分轻松而有趣, 但却击中要害。他指出, 哥德尔的工作并不是反对机械论的论证, 而只是说明了一个看来已深深地暗藏在意识阴谋之中的基本圈套。

18 第七次远足或特鲁尔的徒然自我完善

斯坦尼斯瓦夫·莱姆

宇宙无限却有界，因此，一束光不管它射向哪一个方向，在亿万年之后，将会回到——假如这光足够强有力——它的出发点。谣言也同样，从一个星球到另一个星球，传遍每一处。有一天，特鲁尔听远处的人说，有两个力大无比的建造者兼捐助人，聪明过人，多才多艺，谁也不是他们的对手。他赶忙跑去见克拉鲍修斯。后者向他解释说，这两个人并不是什么神秘的敌人，而正是他们自己，因为他们已经遐迩闻名。然而，名声有一个缺点，即它对人的失败只字不提，尽管这些失败正是极度完美的产物。谁若是不信，就请回忆一下特鲁尔七次远足的最后一次，那次他没与克拉鲍修斯结伴同行，后者因有要事而不能脱身。

在那些日子里，特鲁尔非常自负，他接受了各种各样应得的崇誉和称号，这都是十分正常的。他驾着飞船向北飞去，由于他对这个区域不熟悉，飞船在渺无人烟的空间里航行了好一段时间，途中经过了充满战乱的区域，也经过了一些现已变得荒芜

寂静的区域。突然，他看见了一颗小星球，与其说这是一颗星球，倒不如说是一块流失的物质。

就在这块大岩石上，有人在来回奔跑，奇怪地跳着脚，挥着手。对这个无比孤独、绝望、也许还是愤怒的人，特鲁尔感到惊讶，也感到关切，于是，他立即把飞船降落了。那个人就向特鲁尔走来。此人显得异常傲慢，浑身上下都是铁和钒，发出叮呤当啷的金属碰撞声。他自我介绍说，他是鞑靼人埃克塞尔修斯，曾是潘克里翁和西斯班德罗拉两大王国的统治者。这两个王国的臣民一时疯狂而将他赶下王位，放逐到这颗荒芜的小星球上。从此，他便永远在黑暗和流星群中飘游。

当这位被废黜的国王知道了特鲁尔的身份后，就一个劲地要求他帮助自己马上恢复王位，因为特鲁尔做起好事来也是个专家。那位国王想到复位，眼中燃烧着复仇的火焰，他那双高举的铁手紧握着，仿佛已经掐住了那些可爱的臣民的脖子。

特鲁尔并不想按照国王的要求行事，因为这样做会造成极大的罪恶和痛苦，但他又想安慰一下这位蒙受耻辱的国王。思索片刻之后，他觉得事情还有补救的希望，因为完全满足国王的心愿还是可能的——而且不会让那些百姓遭殃。想到这里，他卷起衣袖，施展出他的全副本领，给国王变出了一个崭新的王国。新王国里有许多城市、河流、山脉、森林和小溪；天空中飘着白云；军队骁勇无比；还有许多城堡、要塞和淑女的闺房；繁华的集市在阳光下喧嚣不止，人们在白天拚命干活，到了晚上则尽情歌舞到天明，男人们还以舞刀弄剑为乐。特鲁尔想得很细，还在这个王国里放进了一座用大理石和雪花石膏建造的豪华首都。在这里，聚集着一群头发灰白的贤人；还配有过冬的行宫和消夏的别墅；这里也充斥着阴谋家、密谋者、伪证人和告密者；大路上奔驰着浩浩荡荡的骑兵队伍，红色的羽毛饰

迎风招展。特鲁尔还别出心裁，使嘹亮的号声划破天空，紧接着是二十一响的礼炮声，他还往这个新王国里扔进一小撮叛国者和一小撮忠臣，一些预言家和先知，以及一个救世主和一个伟大的诗人。做完这些之后，他弯下腰，发动起机关，并用微型工具作了最后的调整。他给那个王国的妇女以美貌，给男人以沉默与酒后的粗暴，给官吏以傲慢与媚骨，给天文学家以探索星球的热忱，给孩子们以擅长吵闹的能力。所有的这些都被特鲁尔有条不紊地装进一个盒子里，盒子不太大，可以随身携带。他把这个盒子赠给可怜的国王，让他对它享有永久的统治权。他先向国王介绍了这个崭新王国输入和输出的所在，教他怎样编制关于战争、镇压暴乱、征税纳贡的程序，还向他指明了这个微型社会的几个关键之处，哪些地方最易发生宫廷政变和革命，哪些地方则最少有这类变动。特鲁尔把一切有关的情况都作了仔细的介绍，而国王又是统治王朝的老手，马上就领会了一切，于是，在特鲁尔的监督下，他试着发布了几个号令，他准确地操纵着控制杆，控制杆上面雕刻着雄鹰和勇狮。这些号令一宣布，全国便处于紧急状态，实行军事管制和宵禁，并对全体国民征收特别税。王国里的时间过去了1年，而对在外面的特鲁尔和国王来说，还不到1分钟。国王出于功德无量的宽容，也就是说用他的手指在控制杆上轻轻地拨了一下，便豁免了一个死刑犯，减轻了特别税，撤销了紧急状态，于是，全体臣民齐声称谢，欢呼声就如同小老鼠被倒提尾巴时发出的尖叫。透过刻有花纹的玻璃，你可以看到，在尘土飞扬的大道上，在水流缓慢的河岸边，人们在狂欢，齐声歌颂统治者的大恩大德。

由于盒子里的王国太小，像小孩的玩具，起先这位国王还颇不满意，但是，当他透过盒子的厚玻璃顶盖看去，发现盒中的一切看上去都很大时，他慢慢地有所领悟，大小在此无关宏

旨，因为政府是不能用公尺和公斤来衡量的，感情也同样，无论是巨人还是侏儒，他们的感情很难有高矮之分。因此，他感谢了制造这个盒子的特鲁尔，即使态度多少有点生硬。又有谁会知道这位狠毒的国王在想些什么呢？也许此刻他正在肚中盘算着将他的恩人特鲁尔套上枷锁，折磨至死，杀人灭口，免得以后有人说闲话，说这位国王的王朝只不过是某个以四海为家的补锅匠的微薄施舍。

然而，由于他们大小悬殊，这位国王很明智，认为这是绝不可能的，因为还没等他的士兵抓住特鲁尔，后者敢几个跳蚤便可将他们统统抓住。于是，他又一次冷淡地向特鲁尔点了一下头，把象征王权的节杖和圆球夹在腋下，双手捧起盒子王国，咕哝一声，走向那流放时住的小屋。外界，炽热的白昼与混沌的黑夜交替着，被臣民们认为是世上最伟大的国王根据这颗小行星的旋转节奏，日理万机，下达各种手谕，有斩首，也有奖励，使得百姓对他忠心耿耿，百依百顺。

特鲁尔回到了家中，不无自豪地将这件事告诉了克拉鲍修斯。他将事情的经过一一讲出，说起他如何略施小计，既满足了国王的独裁欲望，又保障了他以前的臣民们的民主愿望，言谈间不禁流露出得意之情。但令他吃惊的是克拉鲍修斯并没有赞赏他，相反，他脸上显出责难之色。

沉默片刻之后，克拉鲍修斯终于开口了："你是不是说，你把一个文明社会的永久统治权给了那个杀人不眨眼的暴君，那个天生的奴隶主，那个以他人的痛苦取乐的虐待狂？而且，你还对我说他废除了几个残酷的法令便赢来了一片欢呼声！特鲁尔，你怎能做出这样的事？"

"你这是在开玩笑吧？"特鲁尔大声说道。"事实上，这个盒子王国才 2 英寸长，2 英寸宽，2.5英寸高……这不过是个模

型……”

“什么东西的模型？”

“什么东西？当然是一个文明社会的模型，只不过缩小了几亿倍。”

“既然如此，你又怎么知道天下没有比我们大几亿倍的文明社会？如果真有的话，我们这个文明社会不就成了模型了？大与小有什么关系？在盒子王国中，居民们从首都去边远省份不也要花几个月的时间吗？他们不也有痛苦，也有劳累，也会死亡吗？”

“请等一下，你很清楚，所有这些过程都是根据我设计的程序进行的，因此它们不是真的……”

“不是真的？你的意思是说盒子是空的，里面发生的游行、暴力和屠杀都是幻觉？”

“不，不是幻觉，因为它们具有实在性，只是这种实在性完全是我通过摆弄原子而导致的微型现象，”特鲁尔分辩说。“问题的关键在于，那里发生的生生死死、恩恩怨怨，只不过是电子在空间的轻微跳跃，完全听从我的非线性工艺技术的安排，我的技术……”

“行了行了，别再吹了！”克拉鲍修斯打断了他。“那些过程是不是自组的？”

“当然！”

“它们是在无穷小的电荷云中发生的？”

“你知道得很清楚，当然是的。”

“那么，那里发生的黎明、黄昏、血腥的战争都是因真实变量的相互作用而产生的？”

“正是的。”

“如果你用物理、机械、统计和微观的方法来观察我们这

个世界，不也是些电荷云的轻微跳跃吗？不也是正负电荷在空间的排列吗？我们的存在不也是亚原子的碰撞和粒子的相互作用的结果吗？尽管我们自己把这些分子的翻转感知为恐惧、渴求或静思。当你在白日里遐想时，在你大脑里除了相联与不相联环路的二进制代数和电子的不断游动外，还有什么呢？”

“你说什么，克拉鲍修斯？难道你认为我们的存在与那个玻璃盒里的模拟王国是一样的？”特鲁尔慷慨陈词。“不，不一样，这完全是风马牛不相及的！我的目的只不过想制造一个国家的模型，这个模型只从控制论的角度来看是完美的，仅此而已！”

“特鲁尔！我们的完美正是我们的灾难，因为我们每前进一步，都将招致无法预料的后果！”克拉鲍修斯的声音越来越大。“如果一个拙劣的模拟者想要折磨他人，会制造一个木偶和蜡像，然后使它大概有个人样，这样，不管他怎样拳打脚踢，也完全是微不足道的讽刺而已。但如果这场游戏有了一系列的改进，情况就会大不一样。比方说，有这样一个雕塑家，在他的塑像的肚中安装了一个放音装置，只要照准它的腹部打去，它就会惨叫一声。再比方说，要是一个玩偶挨了打会求饶，就不再是个粗糙的玩偶了，而是一个自稳态生物；如果一个玩偶会哭，会流血，知道怕死，也知道渴望安宁的生活，尽管这种安宁只有死亡才能带来！你难道看不出来，一旦模拟者完美无缺，那么，模拟和伪装就变成真事，假戏就会真做！特鲁尔，你让多少个血肉之躯在一个残酷的暴君手下永远受折磨……特鲁尔，你犯下了一个弥天大罪！”

“这纯属诡辩！”特鲁尔厉声喊道，因为他此刻已感到了他朋友话中的含义。“电子不仅在我们的大脑里游动，它们同样也在唱片中游动，这并不能说明什么问题，当然也不能证明你这种类推！那个魔鬼国王手下的百姓们被砍了头也确实会死，也

知道信心、战斗，还会相爱，因为我建立的参数正是这样。但是，克拉鲍修斯，你不能说他们在这个过程中会有什么感觉，因为在他们大脑中跳跃的电子不会告诉你这方面的知觉！”

“但是，如果我窥视你的大脑，也只能看到电子，”克拉鲍修斯反驳说。“好了，不要装傻了，别假装不明白我的意思了，我知道你不至于那样蠢！你想想，一张唱片会听你差遣，会跪地求饶吗？你说你无法分辨那些臣民挨了打之后是真哭还是假哭，因为你不知道他们是因为电子在身内跳跃而发出车轮碾地那样的尖叫，还是因为真的感觉到了疼痛而失声痛哭。这个区别好像很有道理。但是，特鲁尔，痛苦是看不见、摸不着的，只要一个人的行为有痛苦的表现，那他就是感觉到了痛苦！你此时此刻请拿出证据给我一劳永逸地证明，他们没有感觉，没有思维，没有意识到他们在生前与死后之间的这段空白。特鲁尔，你把证据拿给我看看，我就算服了你！你把证据拿出来，证明你只**模拟**了痛苦，而并没有**创造**痛苦！”

“你心里太清楚了，这是不可能做到的，”特鲁尔平静地答道。“即使当盒子里还一无所有，我还没有拿起工具的时候，我就预料到有这样一种求证的可能性，我的目的是为了消除这种可能性。不然，那个国王迟早会发现他的臣民不是真人，而是一群傀儡，一群木偶。你应该理解，没有其他的办法！一旦让国王发现有半点蛛丝马迹，那就会前功尽弃，整个模拟就会变成一场机械游戏……”

“我明白，我太明白了！”克拉鲍修斯大声说道。“你有崇高的愿望，你只想建造一座能以假乱真的王国，由于鬼斧神工，没有人能够辨出真假，我认为，在这一方面你成功了！你虽然回来了才几个小时，但是对于那些被囚禁在盒子里的人们来说，几百年的光阴已经流逝了，有多少生灵遭到蹂躏，而这纯粹是

为了满足那个国王的虚荣心！”

听到这里，特鲁尔二话没说，拔腿就向他的飞船跑去，但他发现他的朋友也紧随其后。特鲁尔的飞船直驶空间，开足马力，朝远处两大团火光之间的那条虹飞去。在路上，克拉鲍修斯对他说：“特鲁尔，你真是不可救药。你做事总不三思而行。到了那儿之后，你打算怎么办呢？”

“我要把那个王国从国王的手里夺回来！”

“夺回来以后又怎么处置呢？”

“毁了它！”还没等话说完，特鲁尔意识到自己这话的意思，赶紧住了口。最后，他喃喃地说道：

“我要举行一次选举，让百姓们从他们自己中间选举出公正的领袖。”

“你的程序把他们设计成为封建君主的唯命是从的顺民，选举又能解决什么问题？首先，你必须砸碎整个王国的结构，然后从头建立起一个新秩序……”

“但是，从何处结束对社会结构的改变和开始对人的灵魂的改造呢？”特鲁尔大声问道。对此，克拉鲍修斯无以作答。他们相对无言，默默地在宇宙中飞行着，终于看见了那颗小行星。他们先绕着它飞了一圈，准备降落，这时，他们看到了一幅令人惊奇的景象。

整个星球上布满了有智能生命的迹象。每条小河和溪流上都架着微型的小桥，像一条条细线。那些小湖就像水坑，湖水倒映着灿烂的星空，湖面上微小的白帆点点，恰似飘浮的落叶……星球上是黑夜的那一半区域布满了灯光闪烁的城市，而在白昼的那一半区域，隐约可见繁华的大都市，当然，城市里的居民由于太小而无法见到，用高倍望远镜也看不见。特鲁尔根本没有发现国王的踪迹，好像尘土已经把他给淹没了。

“他不在，”特鲁尔吃惊地低声说道。“他们把他怎么样了？他们不知想了什么办法冲出了盒子，占据了整个星球……”

“你看！”克拉鲍修斯说，手指向一朵比顶针大不了多少、状似蘑菇的小小的云彩，它慢慢地升向天空。“他们已经发现了原子能……你瞧那边，你看那块小玻璃。这是你那个盒子的残骸，他们将它供在神庙里……”

“我不明白，这毕竟是一个模型呀。这只是一个具有大量参数的过程，一种模拟，一个假设的王国，里面有必需的反馈、变量等东西……”特鲁尔目瞪口呆，喃喃自语。

“是这样。但你犯了一个不可饶恕的错误。你把你的复制品弄得太完美无缺了。你不想将它们制作得像钟表那样死板，因此你精心设计，无意中创造了一件可能的、合乎逻辑的、必然的东西，而这件东西又变成了机械论的对立面……”

“求求你，别再说了！”特鲁尔说道。他们俩默默地凝视着那颗行星。突然，有样东西撞了一下他们的飞船，实际上是轻轻地擦了一下。他们借着那个物体尾部发出的火带般的光亮，看清了它的面目。这可能是条飞船，也许是颗人造卫星，但实际上与那个国王以前脚上穿过的铁靴很相似。他们抬眼望去，看见那个小星球的上空闪烁着一个天体，特鲁尔以前没有见过它。他们先看到了那根白色的象征王权的节杖，这不是国王的手中之物吗？接着，他们又辨认出国王那冷漠的脸。这位国王就是这样变成了这个小人国上空的月亮。

D·R·霍夫施塔特

D·C·丹尼特

反　　思

妇人既然常啼哭　　心中必定有悲伤

——安德鲁·马弗尔

“但是，特鲁尔，痛苦是看不见、摸不着的，只要一个人的行为有痛苦的表现，那他就是感觉到了痛苦！”

莱姆在描写他那奇特的模拟时精心的措词，非常有意思。在他的故事里，像“数字”、“非线性”、“反馈”、“自组”和“控制论的”这样的字眼反复出现。因此，它们别有一种不赶时髦的韵味，与目前对人工智能的讨论所使用的大部分名词截然不同。人工智能的许多研究工作已误入迷途，与感知、学习和创造性相去甚远，更多的是讨论如何模拟运用语言的能力等方面(“模拟”这词是我们的提法)。我们认为，在人工智能研究中，难度更大、争议更大的课题还未曾触及，人心的“自组”和“非线性”的本质将来会作为探讨的重点继而重新出现。与此同时，莱姆生动地运用了那些本来就很有魅力的词汇。

汤姆·鲁宾逊有一部小说，题为《连牛妹也会伤心》。在这部小说里，有一段描写与莱姆的微型世界有惊人的相似之处：

那年圣诞节，朱利安把一个蒂罗利昂村的小模型作为礼物送给西西，这个小模型十分精致。

村里有一座教堂，阳光照在彩色玻璃上，呈现出水果色拉般的色彩。还有一个广场和一个屋外花园酒店。每逢周六晚上，花园酒店便格外热闹。面包坊永远散发着烤面包的浓香。市政厅和警察局机构庞大，官僚主义和腐化堕落一目了然。个头较小的蒂罗利昂人穿着用皮革精心缝制的衣服，皮衣里面的内衣做工也很讲究。滑雪板商店临街而立，还有其他许多有趣的设施，其中包括一座孤儿院。这座孤儿院就是为了在圣诞之夜引来点火焚烧的，每年都要被烧掉。孤儿们穿着着火的睡衣从火堆中窜出来，在雪地上打滚。真可怕。大约在一月的第二个星期，来了一个火警检查员，用一根木棒在灰烬中拨来拨去，口中还念念有词：“要是这些孩子早点听我的话，就不会葬身火海了。”

这段描写尽管与莱姆的故事主题相同，但情趣各异。就好像两位作曲家各自想到了同一个旋律，但又各有乐思。鲁宾逊让你觉得这个小人村简直是不可思议的精致的机械作品（如果你不是愚蠢得出奇的话），绝对没让你去相信这些小人有真情实感。

年复一年的火烧孤儿院的游戏，充分反映了尼采的永恒轮回观，即世上发生的万事万物会反复重演，这个事实使得这个小人世界失去了任何真实意义。为什么那位火警检查员对殉难儿童的哀悼那么空洞呢？是蒂罗利昂人自己反复制造孤儿呢，还是有一个“复建”的按钮？源源不断的孤儿从何而来？会不会是火中的孤儿又“死”而复“生”？至于这里面其他一些奇思妙想，细读一下已省略的细节会受益匪浅。

你是否因此相信这些小心灵确有其事，这完全取决于作者妙笔生花的文采和叙述的技巧。你相信哪一种说法？

19 心灵、大脑和程序

约翰·R·塞尔

近年来，计算机对人类的认知能力进行了模拟，这种模拟的心理学和哲学意义是什么呢？为了回答这个问题，我觉得有必要先区分几个不同的概念，它们是“强的”、“弱的”和“谨慎的”人工智能。根据弱人工智能这个概念，计算机在研究心灵方面的主要价值是提供了十分得力的工具。比方说，有了计算机，我们能够更严密、更精确地提出假设，并且对之进行测试。但是，根据强人工智能这个概念，不仅可以说，计算机是研究心灵的工具，而且具备恰当程序的计算机本身就是心灵。因为一旦计算机有了正确的程序，我们就能说，它实际上就有了理解力以及其他的认识状态。在强人工智能里，由于输入了程序的计算机具有认识状态，因此，程序不仅仅只是帮助我们测试心理学解释的工具，程序本身就是解释。

我对弱人工智能的看法并没有反对意见，至少在本文是如此。我的重点将放在属于强人工智能观点的一些看法上，特别是其中的一种意见，即认为，配备适当程序的计算机事实上确

有认识状态，并且能借此解释人类的认知。后面当我提到人工智能时，我指的是用上述两种说法表述的强人工智能的概念。

我将研究的是罗杰·香克和他在耶鲁大学的同事们所做的工作(请参阅香克和艾贝尔森1977年的著作)，因为比起其他类似的观点来，我更熟悉他们的看法，而且，他们的看法为我要研究的问题提供了十分清晰的例子。但是，问题的讨论又不依赖于香克所设计的程序的细节，同样的论证也适用于威诺格拉德的SHRDLU程序(威诺格拉德，1973年)。魏岑鲍姆的ELIZA程序(魏岑鲍姆，1965年)其实适用于任何对人类心灵现象所作的图林机模拟。

如果省略各种细节，我们可以用三言两语对香克的程序作如下描述：这个程序的目的是为了模拟人类理解故事的能力。人类理解故事的能力有个特点，尽管故事对有些事情并没有加以明确说明，但是，人们还是能够回答与其有关的问题。比方说，你听到这样一个故事："有人走进一个餐馆，要了一份汉堡牛排。菜上来了，那人一看，只见盘中是一块被炸得黑乎乎的牛排，于是，他大发雷霆，拂袖而去，既没结帐也没留下小费。"听完这段故事，如果有人问："那个人吃没吃那份汉堡牛排?"你也许会回答说："没有，他没吃。"同样，如果你听到这样一个故事："有人走进一家餐馆，要了一份汉堡牛排。菜上来了，那人一看，十分高兴。他给了女服务员一笔可观的小费，并结了帐，然后离开了这家餐厅。"现在，要是问："那个人吃没吃那份汉堡牛排?"你可能会说，"是的，他吃了。"香克的计算机也能够以同样方式回答有关餐馆的类似问题。为了做到这一点，计算机应对人类关于餐馆的那种知识，有一种"表象"，以便在听了这样的故事以后，能够回答类似于上面的问题。在计算机听完故事回答问题时，所给出的答案，将与人类听完类似

的故事之后所给出的答案应该是一样的。鼓吹强人工智能的人认为，在这个问答过程中，计算机不仅在模拟人类的一种能力，而且，(1)实际上是理解了故事并对问题作出了回答。(2)计算机及其程序所做的一切的确解释了人类理解故事并对有关问题作出回答的能力。

在我看来，香克的研究工作根本没有对这两种看法提供依据，我将在下面进行说明(当然，我并不认为香克本人也持这种看法)。

测试任何心灵理论正确与否的方法之一是问你自己，如果你的心灵真的按照那个理论所说的普遍原则活动的话，那将会是一种什么样的情形？我们用下列的思维实验来这样测试香克的程序。假设我被锁在一间屋子里，有人给我看许多汉字。再假设(实际上也确实如此)我不谙汉语，既不会说，也不会写，甚至吃不准会不会把汉字与日文或没有意义的涂写相混淆。在我看来，汉字就是没有意义的龙飞凤舞的曲线。现在，再让我们假设，在看过这第一批汉字后，有人又给我看第二批汉字，这一次附有一套将第一批汉字与第二批汉字联系起来的规则，这套规则是用英语写的，我同其他母语是英语的人一样，对这些规则的理解没有任何问题。凭着这些规则我能够将一套形式符号与另一套形式符号联系起来，这里的“形式”意思是说我单凭外形就完全能够分辨出这些符号。现在假设，有人给我看了第三批汉字，并附有一些也是用英语写的指令，这些指令使得我能够将第三批汉字与前两批联系起来。这些指令教我如何用具有某些形状的汉字符号对给我看的第三批汉字中某些形状的汉字作出反应。我根本不认识给我看汉字的人，他们把第一批汉字叫作“手迹”，把第二批汉字叫作“故事”，把第三批汉字叫作“问题”。我对第三批汉字作出的符号反应，被称之为“问题的

回答”，他们给我的用英语写成的规则则被称之为“程序”。现在，我们稍微把这个故事变得复杂一些。让我们设想这些人给我看的故事是用英语写成的，英语我当然是很熟悉的了，然后，他们用英语向我提问题，我则用英语作出回答。再假定，过不了多久，我已经能够十分熟练地根据那些指令来使用汉字符号了，程序的设计者也能够信手写出程序来，使得从外部的观点看来，即从我的屋外的某个人看来，我作出的回答与母语是汉语的人所作的回答根本就难辨真伪。如果光是看我的回答，根本就无法知道我对汉语一窍不通。再假设，其实事实上也是如此，我对用英语的提问所作的回答与其他母语是英语的人所作的回答丝毫没有两样，原因很简单，因为我的母语也是英语。从外部的观点看来，即从某个只是看见我的“答案”的人的观点看来，对汉语提问与对英语提问的回答一样出色。可是，在回答汉语提问时与回答英语提问时不同，我是通过使用一些未经解释的形式符号来提供答案的。就汉语而言，我的行为纯粹像一台计算机，我只能对规定形式的单元进行计算机运算。如果单从汉语的角度看，我只不过是计算机程序的一个例示。

强人工智能的观点认为，具有程序的计算机能够理解故事，并且，从某种意义上说，这种程序解释了人类的理解能力。现在，我们能够通过我们的思维实验来检验这两种说法是否正确。

(1)关于第一种说法，我认为，通过上面所举的例子，我们可以清楚地看到，我对用汉语写成的故事一个词都不理解，尽管我所有的输入和输出与母语是汉语的人简直难辨真伪，我用什么样的形式程序都可以，但我实际上什么都不懂。出于同样的理由，香克的程序对故事的内容一点也不理解，不管是用汉语写成的，还是用英语写成的，还是用其他语言写成的，都无

所谓，因为在这个汉语的例子里，我就是计算机，而在那些我不是计算机的情况下，计算机同我不识汉字的情况一样，什么都不懂。

（2）关于第二种说法，即程序解释了人类理解力，我们可以看到，计算机及其程序并没有提供充足的理解条件，因为计算机只是在操作，并未在理解。但是，计算机是不是提供了一个必要的理解条件，或对理解作出了有意义的贡献？强人工智能的支持者们有这样一种看法，当我理解了一个英语故事时，我所做的一切与我使用汉字符号时所做的一切没有什么两样，或许是相差不多。我懂英语的情形与我不懂汉语的情形区别只在于对形式符号的使用。我并没有论证这种看法是错误的，但毫无疑问，这种看法在上述例子里是不可信的。这种看法的貌似合理性来自这样一个假设：我们能够设计一个程序，这个程序具有像操母语的人一样的输入和输出，另外，我们还假定说话者具有某种描述层次，但在这种层次上，他们仍然是程序的例示。基于这两个假定，我们说，即使香克的程序没有完全解释理解能力，但也许作出了部分的解释。我假定这是一种经验的可能性，但它并没有提供丝毫证据，使我们相信这种可能性。这是因为这个例子表明（尽管肯定没有论证）计算机程序与我对那个故事的理解毫不相关。当我看汉字时，脑袋里的东西都是人工智能通过程序输入的，我对那些汉字根本不理解。当我看英文时什么都理解，这就无需假定我的理解会与计算机程序有什么相关，也就是说，会与对纯规定形式的单元作计算机运算有什么相关。只要程序被定义为是对纯规定形式的单元所进行的计算机运算的话，那么，这个例子说明，这些运算本身与理解根本就没有什么有意思的联系。它们根本就不构成理解的充分条件，也丝毫没有理由假定它们构成了理解的必要条件，甚

至说它们对理解作出了重大的贡献。请注意，这个论证的力量不仅在于，根据不同的形式原则进行运算的不同计算机能够具有同样的输入和输出，这根本不是问题的关键。此论点的关键在于，不管你在程序中存入何种纯形式的原则，它们都不能为理解提供充分的条件，因为一个人能够在不理解任何东西的情况下，采用这些形式原则。根本就没有任何理由认为，这些原则是充分的并具有重大贡献的，因为没有任何理由说，当我理解英语时，我其实是在用某种形式程序进行运算。

那么，在我读英语句子与我读汉语句子之间又有什么不同之处呢？显而易见的答案是，我明白英语句子的意思，而我却完全不明白汉语句子的意思。但这个区别又在哪里呢？不管它是什么，我们为什么不把它给计算机呢？这个问题我将留在后面进行讨论，现在我想继续解释这个例子。

我曾有机会将这个例子讲给几位人工智能的研究者听，有趣的是，对什么是对它的恰当回答，他们似乎意见并不一致。各种各样的回答使我感到不胜意外。下面我想举几种最普通的回答(并且，我将注明这些回答的地理位置)。

但是，我想先纠正对“理解”这个概念的一些误解。在许多这样的讨论中，你会发现“理解”这个词有许多相当策略的用法。批评我们的人指出，理解具有各种不同的程度；“理解”并非一个简单的双位谓词；它具有不同的种类和层次，连逻辑中的排中律也常常不能直接适用于像“X理解了Y”这样的陈述；另外，在许多情况下，X是否理解了Y这个问题纯属判断，并非事实就是如此。对于所有这些看法，我只想说：当然，当然。但是，它们与我们正在讨论的问题无关。有些明显的例子说明，“理解”的确适用，也有些明显的例子说明它并不适用。我的这个论证就是需要这两种例子。① 我完全理解用英语写的故事，但理解

用法语写的故事就差一点，用德语就更差，用汉语则一窍不通了。但是，我的汽车和计算器没有丝毫理解可言，它们与此事没有任何关系。我们常常通过比喻与类推，认为汽车、计算器以及其他人工制品也能“理解”和进行其他认知活动，但这种做法什么也没有证明。我们说“这扇门因为装有光电池而知道什么时候开”，“计算器知道如何（懂得如何，能够）做加法和减法，但不会做除法”，还有“恒温器能够感知湿度的变化”。我们说这些话的原因十分有意思，它必然与这样一个事实有关，即我们将自己的目的性延伸到人工制品之中。② 我们所制造的工具是我们目的的延伸，因此，我们很自然地借助于比喻认为它们具有目的性。

但是，我认为这例子并没有说明什么哲学问题。一扇自动门通过光电池“理解指令”与我们理解英语的意思完全不同。如果认为香克的带程序计算机理解故事就同自动门理解指令一样，而不是同我理解英语一样，那么，这个问题就没有必要再进行讨论了。但是，纽厄尔和西蒙(1963年)在其著作中写道，计算机的认知与人类的完全一样。我喜欢他们的直率。我要研究的正是这种看法。我想要论证，带程序的计算机理解的实际上就是汽车、计算器理解的东西，也就是说，它们其实谈不上理解了什么。计算机的理解不是部分的或是不完善的（就像我对德语理解一样），而是等于零。

现在让我们看一看各处的人对这个问题的不同回答：

---

① 还有，“理解”既暗示了拥有心理（有目的的）状态，又暗示了拥有这些状态的真值（有效性和成功）。出于对本文讨论的考虑，我们只研究拥有这些状态时的情况。

② 根据定义，目的性是某些心灵状态的特征，这些心灵状态是针对有关外部世界对象或事态的。因此，信仰、欲望以及目的属于有目的性的状态，而无目标的焦虑和沮丧则不在其列。

(1) 系统回答（伯克利分校）。 "一个被锁在房间里的人的确不理解故事的内容；但事实上这个人只是整个系统的一部分，而这个系统确实理解故事的内容。在这个人的面前放着一大张分类表，表上尽是些规则，他还有许多纸和笔，可以进行计算，另外，他还有汉字符号的'数据库'。现在，理解不是纯粹个人的事，而属于整个系统，个人只是系统的一部分。"

对于这种系统理论，我的回答十分简单：让这个人将该系统的所有部分都内在化。换句话说，让他记住分类表上的规则和汉字符号的数据库，并且，通过心算来代替笔算。这样，个人就体现了整个系统。系统能做的事情，个人也完全能够做到。我们甚至能够抛弃锁在房里的前提，而设想他在户外工作。同样道理，如果个人不理解汉语，那么，系统就更不理解了，因为系统所有的一切没有不装入他的脑子里的。如果个人不理解，那么系统就更没有办法理解了，因为系统只不过是个人的一部分。

实际上，我对系统理论作出这样的回答，甚至都觉得有点不好意思。因为这个理论在我看来实在是不屑一顾的。那些人认为，个人不理解汉语，然而，个人与若干张纸加在一起居然就能无师自通，这真是咄咄怪事。很难想象，一个不信某种教条的人会认为这种理论是有道理的。然而，我想许多相信强人工智能论的人最终会发表类似这种理论的言论，因此，索性让我们再深入一点。这个理论有这样一种说法，它认为，尽管在内在化系统例子中的那个人不像母语是汉语的人那样理解汉语（因为，比方说，他不知道此故事讲的是餐馆和汉堡牛排等事情），但是，"作为形式符号操作系统的那个人"真的理解汉语。处理汉语形式符号的子系统不该与处理英语形式符号的子系统

相混淆。

因而，实际上在这个人身上有两个子系统：一个懂英语，另一个懂汉语，而且，“这两个系统互不相干”。但是，我想指出的是，它们不仅互不相干，而且甚至有天壤之别。懂英语的那个子系统(允许我们暂时用这样的术语来说话)知道那个故事讲的是餐馆和吃汉堡牛排，这个人知道所提的问题与餐馆有关，他也知道他正尽量根据故事的内容作出各种推断来回答问题，如此等等。但那个汉语子系统却不知道这些。英语子系统知道“汉堡牛排”这个词指的是一块美味的肉，而汉语子系统只知道“横直”和“撇捺”。根据这个系统，这个人只知道从一端输入了各种形式符号，然后，按照用英语写成的规则对这些符号进行处理，再从另一端输出另一些符号。原来的那个例子主要是为了论证，这种自动符号处理，不论从何种实际意义上说，都不足以理解汉语，因为，尽管那个人能够在“横直”后面写上“撇捺”。但对汉语还是一字不解。要是假定那个人身上有好几个子系统，也无济于事，因为这些子系统比那个人原来的好不了多少，它们与操英语的人(或子系统)没有丝毫哪怕是最遥远的相似点。事实上，我们从上面的例子可以看出，汉语子系统不过是英语子系统的一部分，它完全是根据用英语写成的规则来对符号作没有意义的处理的。

请让我们自问：首先，是什么动机促成了系统回答？换句话说，认为在那个人身上有一个子系统，此子系统实际上理解用中文写成的故事，这个观点其独立的根据何在？就我们所知，唯一的根据是，在例子里，我有与操汉语的说话者同样的输出和输入，我还有一个来回转换的程序。但是，这些例子总的目的是想说明，这些东西还不足以构成像我理解英语故事那样的理解能力。这是因为，一个人，乃至构成一个人的整套系统，

虽然能够具有输入、输出和程序的正确组合，但还不能像我理解英语那样真正理解任何东西。因此，认为在我身上肯定有一个能够理解汉语的子系统的唯一理由就是我有一个程序，而且我能够通过图林测试，能够以假乱真，使人相信我的母语是汉语。但是，关于图林测试本身的可靠性，也正是我们所要讨论的问题之一。这个例子说明，有可能存在两个"系统"，这两个系统都能通过图林测试，但其中只有一个具有理解能力；如果你说既然它们都通过了图林测试，那它们就都有理解能力，这是站不住脚的，因为这种说法无法证明，我身上的那个理解英语的系统比那个只处理汉语的系统更为优越。一句话，系统回答在没有任何论据的情况下，坚持认为这个系统肯定理解汉语，这无非是在以未经证明的假定作为依据来进行辩论。

不仅如此，系统回答还有可能导致另外的荒诞结论。如果我们认为，根据我有某种输入输出以及位于其中的程序这样一个理由，就肯定我有认知能力，那么，似乎所有不具备认知能力的子系统就会都变得具有认知能力了。比方说，在某一描述层次中，我的胃也能进行信息处理，能够实现任何数量的计算机程序，但是，我相信没有人会说，用来消化食物的胃居然也有理解能力(参见彼里欣1980年的著作)。但是，如果我们接受了系统回答，那么，我们就很难不承认胃、心脏、肝脏等等都是具有理解能力的子系统，因为找不到一个根本的方法来区别说汉语子系统能理解与说胃能理解的动机有什么不同。随便提一句，如果说，汉语系统的输入和输出是信息，而胃的输入和输出则是食物和养料，因此两者具有不同之处，这种说法也不在点子上，因为，从行为者的角度以及从我的角度看，不管是食物还是汉字都没有信息，在我看来，汉字只是一些没有意义的横直撇捺。在这个汉字的例子中，只是在程序设计者和翻译的眼中

才能有信息。如果他们想把我的消化器官的输入和输出也当作信息的话，谁也阻拦不了他们。

最后一点，与强人工智能论的一些独立的问题有关，有必要先把主题撇开一会来解释一下。如果强人工智能论是心理学的一个分支的话，那么它就能够辨别哪些是真正的心理系统，哪些不是，它就能够辨别哪些是心理活动的原则，哪些是非心理系统的原则，否则，它就无法告诉我们究竟什么才算是具体的心理活动。心理的与非心理的区别不能只依据表面的观察，而必须依据系统本身的内在特性，否则，任何人都能够随心所欲地把人说成是非心理的事物，而将飓风说成是心理的事物。然而，在强人工智能论的文献里，这个区别常常模糊不清。长此以往，将给强人工智能是一门认知科学的观点带来灾难性的后果。例如，麦卡锡这样写道："象恒温器这样简单的机器也被说成是有信仰的，对于大多数能够进行解决问题活动的机器来说，具有信仰是它们的一大特征"(参见麦卡锡1979年的著作)。任何认为强人工智能论有可能成为心灵理论的人必须深思一下这种说法的言外之意。据说应把下述事实作为强人工智能论的新发现：墙上的一块用来测定温度的金属也有信仰，就跟我们、我们的配偶和孩子有信仰毫无两样。而且，房间里的"大部分"其他机器，像电话、录音机、计算器、电灯开关等，实际上也都有信仰。本文并不是为了反驳麦卡锡的观点，因此，我只强调一下结论，而不作具体的论证。心灵研究的出发点是这样一些事实，例如，人有信仰，而恒温器、电话和计算器则没有。如果你有一个否定这一论点的理论，如果你能提出一个驳斥该理论的反例，那么，这个理论就是错误的。我们觉得，那些对这类情况有论述的人工智能的研究者们认为他们能够摆脱将恒温器看作有信仰的论点，因为他们并没有把它当真看待，而且

觉得任何人都不会将它当真看待的。但我建议诸位，至少仔细想上1分钟，想想必须具备什么才能使墙上的那块金属有那种真正的信仰；有那种旨在适合命题内容和满足条件的信仰；有那种能成为或强或弱的信仰；有那种盲信的或深思熟虑的信仰以及其他所有的信仰。恒温器当不在其列，胃、肝脏、计算器或电话也不在其列。既然我们把这种观点当真看待，那么，它之为真对于强人工智能论声称是一门心灵科学的论点来说就是致命的。现在心灵无处不有。我们想知道的是心灵与恒温器和肝脏之间究竟有什么区别。假如麦卡锡的观点正确的话，那么，人工智能就没有希望告诉我们这个区别了。

（2）机器人的回答(耶鲁大学)。　“假设我们写出一个与香克截然不同的程序。假设我们把一台计算机置入一个机器人的体内，这台计算机不仅将形式符号作为输入而摄入，并将它们作为输出而传出，而且，它会巧妙地操纵机器人，使机器人表现出像感知、行走、移动、敲钉子、进食、饮水等一类行为——想表现什么就表现什么行为。例如，我们可以在这个机器人身上安装一个摄像机，使它有视力，还可以在它身上安装手和腿，使它能够‘活’动，所有这些东西都由它的计算机‘大脑’控制。这样一个机器人将不同于香克的计算机，它具有真正的理解力和其他心理状态。”

关于这个回答，首先值得注意的是它默认了认知不只是形式符号的处理问题，因为它还补充了与外部世界的某些因果关系(参见福多1980年的著作)。对于这种回答，我认为增加一些诸如“知”和“运动”之类的能力具体地说并没有给香克原来的程序增加什么理解能力，泛泛地说，也没有给它增加什么目的性。为了看清楚这一点，请注意上面那个思维实验也适用于机器人

例子。假定不是计算机位于机器人体内，而是我被锁在一间屋子里，也像前面那个阅读汉语故事的例子那样，然后你给我更多的汉字和更多的英语指令，我按照指令将汉字符号彼此匹配，并将它们输出。假设我得到的某些汉字符号来自机器人身上的摄像机，而我输出的其他汉字符号则是用来使机器人内部的发动机运转从而使机器人的手脚活动起来。对此，我自己毫无觉察。需要着重指出的是，我不过是在处理形式符号，而对其他事情一无所知。我从机器人的“感知”装置那里接收到“信息”，然后，我把“指令”发给机器人的运动装置，而我自己对这些事情一点也不知道。我是机器人腹中的小矮人，但与传统的小矮人又不一样，因为我并不知道发生的事情。除了处理符号的规则，其他事情我一概不知。根据这种情况，我认为机器人根本没处在目的状态中，它完全是根据线路和程序而四处移动。还有，在我执行程序的过程中，我也没有处在相应的目的状态中。我只不过是在执行对形式符号进行处理的形式指令。

(3) 大脑模拟器回答（伯克利分校和马萨诸塞理工学院）。

“假设我们设计了一个程序，这个程序并不表现像香克中文字条那样的关于世界的信息，但是，它却模拟了母语是汉语的人在理解汉语故事并回答汉语问题时大脑突触上神经原激活的实际顺序。计算机将汉语故事和有关问题作为输入而摄入，然后，它模拟真正中国人的大脑在处理这些故事时的形式结构，再将汉语答案作为输出而给出。我们甚至还可以这样设想，这台计算机不只有一串程序，而是有一套程序在同时进行运算，就像人脑在处理自然语言时可能出现的那种实际状况一样。在这种情况下，毫无疑问，我们不得不承认计算机理解了故事。如果我们不承认这一点，那不就否认了中国人能理解这些故事么？

在突触的层次上，计算机程序与中国人大脑程序究竟将有或可能会有什么区别呢?”

在反驳这个观点之前，我想扯开去说一句，任何人工智能(或机能主义，等等)的拥护者都会觉得这个观点十分别扭。我认为，强人工智能论的中心论点是，我们没有必要为了知道心灵如何活动而去了解大脑如何活动。据我所知，强人工智能论的基本假设是这样的：有一个对形式元素作计算处理的心理运算层次，它构成了心理活动的实体，它还可以在各种不同的大脑过程中表现出来，就像任何计算机程序都可以通过不同的计算机硬件来实现一样。根据强人工智能论的假设，心灵之于大脑如同程序之于硬件，因此，我们不研究神经生理学也能了解心灵。如果我们必须在了解大脑之后才能了解人工智能，那我们何必还要研究人工智能？然而，即使我们对大脑活动了解得如此清楚，对理解能力的了解却依然雾霭重重。为了使大家理解这一点，我们想象这样一种情形：思维实验中那个只会一门语言的人手里摆弄的不是语言符号，而是一套带阀门的水管。当他得到汉字符号之后，首先查一下用英语写成的程序，看看他应该打开和关闭哪几个阀门。每一根水管相当于中国人大脑中的一个突触，整个水管系统被配备成这样一种状态：当所有恰当的神经原都激活后，也就是说所有正确的阀门都打开后，汉语答案就会从这一串水管的尽头喷涌而出。

那么，在这样一个系统里，理解能力又在哪里呢？整个系统把汉语作为输入而摄入，并在模拟中国人大脑突触的形式结构之后，又把汉语答案输出来。然而，此人肯定不理解汉语，水管也不会懂汉语，但有人却说这个人加上这些水管便会产生理解能力，我觉得这种说法十分荒唐。如果我们相信这种说法，那么，请记住，从原则上讲，此人能够使水管的形式结构内在

化，他也能够凭借想象来进行所有的“神经原激活”活动。大脑模拟器说的一个问题在于模拟了大脑的一些无关痛痒的东西。如果模拟的只是突触上神经原激活顺序的形式结构，那么，就不能模拟大脑的关键，即大脑的因果特性，也就是大脑生产有目的状态的能力。形式特征并不能代表因果特性，这一点已有水管例子证明了：我们可以把形式特征从有关的神经生理的因果特性中删除。

（4）综合回答（伯克利分校和斯坦福大学）。 “上述三种回答单独还构不成对中国人小屋反例的反驳以及对思维实验的反证，但如果你将它们综合起来，那将更使人信服、更有力得多。让我们想象一下，如果装在机器人脑袋里的是一个呈大脑形的计算机，如果计算机的顺序具有人类大脑的所有突触，如果机器人的行动酷似人类的行为，那么，这样一个东西就不只是一台具有输入和输出的计算机了，而应该被看作是一个统一的系统。在这种情况下，毫无疑问，我们不得不认为它是有目的的。”

只要我们对这种东西仅仅知道这么多，那么在这种情况下，我们完全有理由接受这个不容反驳的假设，即这个机器人是有目的的，对此我完全表示同意。说真的，除了外表和行为外，综合回答的其他理由离题太远。如果我们能够制造一个机器人，它的行为许多方面都酷似人类的行为，那么，我们就会认为它具有目的，除非另有原因说它没有。如果是这样的话，我们事先就没有必要知道，这个机器人的计算机大脑只是人类大脑的形式模拟。

但我实在不知道，这个论点对于强人工智能论究竟有多大的帮助。为什么呢？请看：根据强人工智能论，用正确的输入和输出来实现一个形式程序是构成目的性的充分条件，其实是

构成目的性的组成成分。纽厄尔(1979年)曾说过：心理的本质是物质符号系统的活动。但是，在上面这个例子中，我们认为机器人具有目的，这与形式程序毫无关系。我们之认为它有目的是根据这样一个假设，如果机器人的外表和行为酷似我们人类，那么，在没有别的可猜测到的证明的情况下，就应该认为，这个机器人肯定具有与我们人类相同的心理状态，这种状态引起了它们的行为，并通过行为获得了表现，而且，机器人肯定还有一个内在机制用来产生这样的心理状态。假如我们不依靠这样的假设而知道怎样独立地去解释机器人的行为的话，特别是假如我们知道它具有一个形式程序的话，那么，我们就不会认为它是有目的的了。这也正是我反驳前面第二种回答的论点。

假设我们知道，机器人的行为已完全由这样一个事实作了解释，即在机器人内部有一个人，他通过机器人的传感器接收到了未经解释的形式符号，然后又向机器人的运动装置发送了未经解释的形式符号，那么这个人是根据一系列的规则来处理这些符号的。再让我们假设，此人对这些事实一概不知，他只知道对那些无意义的符号该作哪些运算，在这种情况下，我们就会认为这个机器人纯属机械傀儡。如果假定这样一个机械傀儡有一个心灵，那就既没有根据，也没有必要。因为现在已不再有什么理由去认为机器人或者它所属的那个系统是有目的的(当然，在机器人腹中处理符号的那个人除外)。形式符号处理过程进行得很正常，输入和输出也很匹配，但是，目的的真正所在则是那个人，而他对有关目的状态却一无所知，比方说，他看不见机器人的眼睛感受到什么，他不打算移动机器人的胳膊，他也不理解机器人摄入的输入和送出的输出。另外，出于前面已经说明的原因，包括了机器人和人在内的整个系统对这些

事情也一概不知。

为了说明这一点，让我们将这一情况与另外一些情况加以对比，在另外一些情况下，我们发现，我们非常自然地认为像猿、猴这样的灵长目和像狗这样的家畜是有目的的。为什么我们觉得这样做很自然呢？原因大概有二：首先，如果我们不认为这些动物具有目的，就无法解释它们的行为；其次，我们可以看到，动物和我们人类同样有血有肉，有鼻子有眼，还有一张皮毛。如果说动物的行为与假定它的皮下有着类似于人的因果性组织没什么不一致，那我们就既能断定，这个动物的行为背后一定具有心理状态，又能断定这种心理状态一定是其组织与人类相似的机制引起的。除非另有不容许这样做的原因，我们肯定会对机器人作出类似的假定。但是，当我们得知这种行为是某个形式程序所致，并且生理基质的实际因果特性与此无关时，我们就会放弃这一目的性的假设。

对于我们的思维实验的另外两种意见虽然经常可见（因此值得一提），但实在是离题太远了。

（5）他人心灵回答（耶鲁大学）。 “你怎么知道别人理解汉语或其他东西？完全凭他们的行为。既然计算机能够像别人一样（从原则上讲）顺利通过行为测试，那么如果你认为别人具有认知能力，也就必须从原则上认为计算机有认知能力。”

对于这样一种反对意见，实在不必兴师动众，只须三言两语就能解决问题。这里所谈及的问题并不是关于我如何知道别人具有认知状态，而是当认为他们具有认知状态时，他们究竟具有什么样的东西。这个论证的中心在于，他们具有的不仅仅只是计算过程和输出，因为在没有认知状态的情况下，计算过程和输出也能存在。假装麻木不仁不能回答这一论证。在物理科学

里，我们不得不假设物体的实在性和可知性，同样，在“认知科学”里，我们也假设了心理状态的实在性和可知性。

(6)多层回答(伯克利分校)。“你的整个论证假定人工智能只与模拟计算机和数字计算机有关。但这仅仅是因为目前科技才发展到这一水平。你认为因果过程是目的性的关键，(假定你的看法是对的)不管它们是什么，最终我们将制造出具有因果过程的东西来，那将是人工智能。因此，你的论证根本不在于探讨人工智能形成和解释认知状态的能力。”

我对此回答并没有异议，但我想提出一点，这个回答通过把人工智能重新定义成人为地产生和解释认知的东西，而使强人工智能这一工程变得琐碎而无意义。强人工智能论的最初目标在于论证人工智能是一个确切的、明确限定的论题，那种心理过程是对形式限定因素的计算过程。我一直想对这一论题提出异议。如果这一说法经过重新定义不再是原来的那个论题，我的异议也就不成立了，因为对我的异议来说，已不再有一个可适用的可检验的假设。

现在让我们回头去看一看那个我曾答应解答的问题：假定在我原来的例子里，我懂英语而不懂汉语，假定因此计算机既不懂英语也不懂汉语，那么，我身上肯定有某种东西使我懂英语，还缺少某种东西使我不懂汉语。既然如此，为什么不能使计算机也获得这种东西(无论它是什么)呢？

从原则上讲，我不明白为什么我们不能让计算机也具有理解英语或汉语的能力，因为从某种重要的意义上讲，我们的身体和大脑也正是这样的机器。但我确实发现我们有充分的理由不能把这个东西给予计算机。计算机的运算只是被确定为对形式确定的因素所进行的计算过程。换句话说，计算机的运算是某

一程序的例示。我并不因为是计算机程序的例示，才能够理解英语，并具有其他的目的形式的（我想，我是任何数量的计算机程序的例示），而是据我们所知，是因为我是某种具有某个生物（也就是化学和物理）结构的有机体，在某些条件下，这一结构必定能够产生知觉、行为、理解、学习和其他目的现象。在这个论证中，有一部分理由是说，只有当某样东西具备了这些因果力量之后才有可能获得这种目的性。也许其他的物理和化学过程也能够产生同样的结果。比方说，火星人也能够具有目的性，但他们的大脑完全是用另外一种物质构成的。这是个经验的问题，它有点像另一类问题，比方说，其化学成分与叶绿素不同的某样东西是否也能进行光合作用。

但是，这个论证的中心论点在于，纯形式的模式本身不足以产生目的性，因为形式特征本身不是目的性的构成要素，它们本身没有因果力量，只有一种力量，那就是当计算机在运转时，它们能够产生下一阶段的形式程序。这种形式模式具体实现后所呈现的任何其他因果特征都与此形式模式无关，因为我们能够用完全不同的方式来实现这一形式模式，而前面的因果特征就会荡然无存。即使是操汉语的人奇迹般地完全实现了香克的程序，我们还是能够将同一程序给予操英语的人、水管或计算机，尽管他们或它们都不懂汉语，但却懂得程序。

大脑活动的关键并不在于突触顺序的形式影象，而在于此顺序的实际特征。据我所知，强人工智能论的所有论证都旨在给认知的影象划一个轮廓，并声称这影象就是实在。

作为总结，我想把隐匿在此论证中的某些一般的哲学问题阐明一下。为明瞭起见，我将使用问答的形式，第一个就是大家已经谈腻了的问题。

“机器是否能够思维？”

答案很明显：能够思维。我们人类正是这样的机器。

“好，但是人造的机器是否也能够思维？”

假设我们能够制造一台机器，此机器具有神经系统，里面的神经原具有轴突和树突，其他部分也与人类一样，那么，对此问题的回答还是很明显：能够思维。如果你能够准确无误地复制前因，你也就能够准确无误地复制后果。我们确实也有可能复制出意识、目的性和其他东西。为此，我们将使用与人类不同的化学原理。正如我在前面说过的那样，这是个经验的问题。

“好的，但是一台数字计算机是否能够思维？”

如果你这里所说的“数字计算机”具有一个描述层次，在此层次上，它能够被正确地描述为计算机程序的例示，那么，回答当然是：能够思维。因为我们就是任何数量的计算机程序的例示，我们是能够思维的。

“但是，如果单是计算机加上合适的程序，它是否能够思维，能够理解？单是实现一个程序，当然是正确的程序，是不是产生理解能力的充足条件？”

我认为这个问题提得很好，尽管此问题常常与前面的一个或数个问题混为一谈。对此问题的回答是：不能思维。

“为什么不能？”

因为形式符号的处理过程本身并没有目的性，它们是没有意义的；更严格地讲，它们还不是符号处理，因为这些符号并不代表任何事物。用语言学的术语来说它们只有句法而没有语义。计算机具有的那种目的性只是存在于编程序和操作计算机的人脑中，存在于存入输入和破译输出的人脑中。

前面所举的关在小屋内看汉字的例子旨在说明这个问题，一旦当我们将某种东西给予一个确实有目的的系统（一个人），并用这种形式程序来规定他的活动时，你就会发现，这种形式

程序并没有带来额外的目的性。比方说，它丝毫也没有给那个人的理解汉语的能力带来任何帮助。

区分程序与程序的实现是人工智能的一个特点，此特点十分引人瞩目，但是，对于模拟即等于复制的观点来说，它却是致命的。在计算机硬件里对程序与程序的实现加以区分，就相当于对心理活动层次与大脑活动层次加以区分。如果我们能够将心理活动层次描述成一个形式程序，那么，我们似乎就能够在既不对大脑做内省心理分析也不做神经生理分析的情况下，描述心灵的本质。但是，“心灵之于大脑如同程序之于硬件”这个等式却四处受挫，尤其表现在以下三个方面：

首先，对程序与程序的实现加以区别势必会造成这样的后果，即同一个程序可以有各种各样甚至乱七八糟的程序实现，它们根本就没有目的性。魏岑鲍姆(1976年的著作，第二章)就曾详细地解释过如何用一卷手纸或一堆碎石来构造一台计算机。同样，理解汉语故事的程序也能够输入一套水管、风力机或只讲英语的人脑中，但他们或它们并不因此而能理解汉语。首先，诸如石头、手纸、风和水管这样的东西就不可能具有目的性，只有像大脑这样的具有因果力量的东西才能具有目的性，尽管操英语的人有大脑，并能够有目的性，但光凭熟记程序是得不到更多的目的性的，因为熟记程序不能教会他懂汉语。

其次，程序是纯形式的，但目的状态在此意义上却不是形式的。目的状态是以其内容而不是以其形式而被定义的。比方说，天下雨了这个信念并不以它的某种形式状态而定义的，而是以具有满足条件、适用定向等心理内容而被定义的（参见塞尔1979年的著作)。说实话，在这种句法意义上说这样的信念根本就没有形式状态，因为同一个信念在不同的语言系统内可以通过无数不同的句法形式来表达。

第三，我在前面已经讲过，心理状态和事件实质是大脑活动的一种产物，而程序在此意义上却并非计算机的产物。

“如果程序怎么也不可能构成心理过程的话，那为什么还有那么多的人相信程序能构成心理过程呢？至少必须对此作出某种解释。”

我真不知道该如何回答这个问题才好。有人认为计算机模拟就等于被模拟的对象，这种看法一开始就很值得怀疑，因为计算机绝对不仅限于模拟心理活动。没有人会相信计算机模拟的一场特级火灾会把周围的建筑物烧光，也没有人会相信模拟的暴风雨会将我们淋成落汤鸡。究竟是什么原因使人认为计算机模拟的理解力真的能理解事物？有时你会听到这样的说法，要让计算机感觉疼痛或坠入情网极其困难，但疼痛和恋爱与认知或其他任何事情相比根本不分上下。就模拟而言，你只需正确的输入和输出，中间再加上一个程序，将前者变为后者。计算机的一切活动只需这些东西就足够了。不管是关于疼痛、恋爱、认知、大火还是大风雨方面，其错误都在于将模拟与复制混为一谈了。

但为什么会有这么多人曾经觉得——也许仍然坚持认为——人工智能以某种方式复制并解释了心理现象呢？这里有几个原因。我认为如果不把这些原因弄清楚，我们就无法消除人们的错误观念。

首先，也许还是最重要的是，人们对“信息处理”这个概念的理解存在着一些问题。许多研究认知科学的人认为，人类大脑及其心灵进行着所谓“信息处理”的活动；同样，计算机及其程序也在进行信息处理活动。但是，大火和暴风雨则无什么信息处理可言。因此，尽管计算机能够模拟任何过程的形式特征，但它同心灵和大脑却保持着极为特殊的关系，这是因为，一旦

计算机输入适当的程序，最理想的就是像人类大脑这样的程序，那么，大脑和计算机所做的信息处理工作就完全一样，而这种信息处理正是心理活动的本质。但是，这个论证有个问题。它的基础概念“信息”是有歧义的。如果人们在解算术题，或在谈一个故事并回答与此有关的问题时，就是在“处理信息”，那么，从这个意义上讲，程序计算机并没有在进行“信息处理”，它只是在操纵形式符号。程序的设计和输出的破译者用符号来代表世界上的各种事物，这一点计算机是全然没有能力想到的。我再强调一遍，计算机只有句法而没有语义。因此，如果你打入“2＋2＝?”这个问题，计算机会打出“＝4”。但是，计算机全然不知道这个“4”的意思是4，或者它代表了某种意义。这并不是说计算机没有某种第二级信息以便用来破译第一级的符号，实际上，就计算机本身而言，它的第一级符号根本就没有解释。计算机除了符号还是符号。因此，将“信息处理”这个概念用于计算机就会陷入困境：或者将“信息处理”这个概念解释成它意味着目的性是信息处理的一部分；或者不作这样的解释。如果是前者，那么，程序计算机并没有在进行信息处理，而只是在操纵形式符号。如果是后者，那么，尽管计算机是在进行信息处理，但这与计算器、打字机、肠胃、恒温器、暴风雨和飓风所进行的信息处理并没有什么两样，也就是说，它们都有一个描述层次，在此层次上，我们能够将它们描写成一头输入、中间转换、另一头输出信息这样的一个过程。但在这种情况下，是不是将这种输入和输出解释为通常的信息，那就完全取决于外部观察者了。并不能根据信息处理中的相似性来确定计算机与大脑的相似性。

其次，在许多人工智能理论中还残存着行为主义或操纵主义的观点。由于带有适当程序的计算机能够具有类似于人类的

输入—输出模式，因此，我们不禁认为计算机具有类似于人类的心理状态。但是，当我们发现无论从理论上还是从经验上都可能使一个系统在某些领域内具有人类的能力，同时却可以毫无目的性时，我们就应该能够克服这种冲动了。我的台式计算器具有计算能力，但没有目的性。在本文中我一直想说明这样一个问题，即一个系统能具有输入和输出能力，此能力复制了母语是汉语的人的能力，但不管程序是怎样编制的，它对汉语却还是一窍不通。图林测试正是这种恬不知耻的行为主义和操作主义传统的典型。我坚信，如果人工智能的研究者抛弃行为主义和操作主义，那么，关于模拟和复制的许多问题就会迎刃而解。

第三，残存的操作主义又同二元论的残存形式相结合。二元论有这样一个假设，即心灵和大脑不可同日而语；实际上，只有在这一假设下，强人工智能论才有意义。在强人工智能论（以及机能主义）里，程序是至关重要的，它与它在计算机里的实现毫不相干。实际上，就人工智能本身而言，同一个程序可以由一台电子计算机来实现，可以由笛卡儿的心灵实体来实现，也可以由黑格尔的世界精神来实现。我在讨论这些问题的过程中一个极为惊人的发现是，许多人工智能的研究人员对我认为人类实在的心理现象也许依赖于人脑实在的物理—化学特征的观点大为吃惊。但是，如果转念一想，你会发现对此并不应该感到吃惊。因为除非你接受了某种形式的二元论，否则强人工智能工程将寸步难行。此工程是通过设计程序来复制和解释心理现象的，但是，除非心灵不但在理论上而且也在经验上独立于大脑，否则，你就无法进行这项工程，因为程序完全独立于任何具体的实现。除非你相信心灵不仅在理论上而且在经验上也能与大脑相分离——这实际上是一种极端的二元论——否则

你就无望通过设计和使用程序来复制心理现象，因为程序必然独立于大脑或其他任何具体的程序实现。如果心理活动就是对形式符号的计算活动，那么，它与大脑之间就没有什么有意思的联系；唯一的联系可能就是，大脑也是无数可以用来实现程序的机器之一。这种二元论并不是传统的笛卡儿式的二元论，后者认为世界上存在着两种实体。但是从坚持认为心灵的具体心理现象与大脑的实际特征没有内在的联系这种意义上说，这种二元论又是笛卡儿式的。这种潜在的二元论被蒙上了一层面纱，使我们不易察觉。因为人工智能的文献常常痛斥所谓的“二元论”，但这些作者都没有意识到，他们自己的论点恰恰必须以一种极端的二元论为前提。

“机器是否能够思维？”对此，我个人的看法是：只有一种机器能够思维，一种极为特殊的机器，也就是大脑或者具有和大脑一样的因果力量的机器。正是这个原因，关于思维的本质强人工智能论所谈极少，因为关于机器它几乎无话可说。根据它自己的定义，强人工智能论的研究对象是程序，而程序不是机器。不管目的性会是什么东西，有一点是可以肯定的，它是一种生物现象，像哺乳、光合作用等其他的生物现象一样，它的产生与其特有的生物化学结构有着因果性的依赖关系。如果我们用计算机来模拟哺乳和光合作用的形式顺序，就能生产出奶和糖，对此种说法恐怕没有人会相信吧？但是，一涉及心灵，许多人就愿意相信这样一种不可思议的奇迹，其原因就在于根深蒂固、源远流长的二元论：拥护这种理论的人认为心灵是一种形式过程，它跟奶和糖不一样，它完全独立于特殊的物质原因。

为了替这种二元论辩护，人们常常希望把大脑看成一台数字计算机（顺便说一句，早期的计算机常常被称为“电脑”）。但

希望毕竟是希望。当然大脑是一种数字计算机。因为一切都是数字计算机，大脑也不例外。问题的关键在于大脑产生目的性的诱发能力决不可能是单纯地实现一个计算机程序。这是因为，对于任何程序来说，都可能用某样东西来实现它，而依然不具有任何心理状态。不管大脑使用了什么办法来产生目的，它决不可能靠实现某个程序来做到这一点，因为任何一个程序就其本身而言都不足以构成目的性。

D·R·霍夫施塔特

反　思

这篇文章最初收集了各种见解共达28种。许多见解十分精辟，但若将它们都集于本书，未免数量太多，而且有的见解技术性太强。塞尔的文章有许多优点，其中之一就是它通俗易懂。如果你没有受过人工智能、神经病学、哲学或其他与此有关的学科的专门训练，也能读懂他的文章。

我们的观点正好与塞尔相反，但我们发现这位对手十分善辩。我们不打算对他的观点一一作出反驳，我们只想着重研究他提出的若干问题，对他的其他论点的答复将蕴含在本书的其余部分中。

塞尔这篇文章的依据是他那独特的“关在小屋里看汉字的思维实验”。在这个实验里，他怂恿读者去想象有一个人在对付一大堆汉语符号，而这项工作对于一个极为聪明的人工智能程序来说是完全能够胜任的，它可以读用汉语写成的故事，然后用汉语回答与故事有关的问题，其所作所为完全同人一模一样，

足以骗过外部观察者从而通过图林测试。我们认为塞尔在这里犯了一个十分严重的根本性错误，这个测试给人以这样一个印象，即认为人能胜任这样一项工作居然是有意义的。读者由于接受了这个印象，便不知不觉地对智能与符号操作之间的关系形成了一种不可能成立的非现实看法。

塞尔希望将读者引入的这个幻觉(当然，他自己并不以为这是个幻觉!)取决于他能否做到让读者忽略一个十分重要的问题，即存在于不同的概念层次上的两个系统在复杂性上有着天壤之别。一旦他在这一点上成功了，剩下的事他就可以不费吹灰之力了。起初，塞尔邀请读者同他一起手工模拟一个现存的人工智能程序，这个程序能在一定范围内，以一定的方式回答一定种类的问题。若要一个人用手工模拟这样一个程序，或者是其他现有的程序，那就等于要他进行计算机所从事的繁琐的具体计算，这样一来他的工作必定变得艰巨而单调，如果不是花数个星期或数个月，那至少也得花几天的时间。然而，对于这一点塞尔却视若无睹，他像一个老谋深算的魔术师巧妙地把读者的注意力引开了，引向一个能顺利通过图林测试的假设的程序。他已经越过了许多层权限，却连招呼都不打一个。紧接着，他又邀请读者上了他的另一个圈套，跟他去一步一步地进行模拟，去“一同感受缺乏对汉语的理解能力”。这正是塞尔论证的关键所在。

对此，我们基本上倾向于“系统回答”(另外，我们将在后面论证，塞尔的看法在某种程度上实际上也倾向于“系统回答”)。“系统回答”认为：理解能力不属于(碰巧)作为人的模拟者，而应属于整个系统，用塞尔的话来说，该系统还包括“草稿纸”。我们觉得，塞尔的轻率解释暴露了他对事情的真相一无所知。一台具有思维能力的计算机对约翰·塞尔来说简直是无法容忍

的，就像当年盖罗拉莫·萨凯里无意之中发现了非欧几何而不敢承认一样。在十七世纪后期，人们对于新的几何所引起的观念扩张尚不能接受。然而，在过了50年之后，人们又重新发现了非欧几何并逐步接受了它。

也许“人工目的性”如果曾被制造出来的话，也会碰上与非欧几何同样的遭遇。如果真的出现了一个能顺利通过图林测试的程序，看来塞尔不但不会惊叹这个程序所具有的能力，相反，还会顽固地坚持认为它缺少神奇的“大脑因果力量”（不管它们是什么)。泽农·皮利申在反驳塞尔时指出，塞尔的观点十分空洞，他觉得下面这些话使人联想起朱博夫“大脑的故事”（第12章)，他不知道这段话是否能够确切地代表塞尔的观点：

> 如果集成电路块在你大脑中代替了越来越多的细胞，如果每个元件的输入—输出功能与已被替代的脑细胞一模一样，你也许还是照常像眼下一样在谈话，但有一点却不同，你所使用的语言将最终失去原有的意义。在我们外部观察者看来是语言，而在你看来只不过是集成电路块使你发出的一些响声而已。

塞尔观点的不足之处在于他没有提供一个明确的方法，用以确定在什么时候，真正的意义——实际上就是真正的“你”——从系统中消失了。他只是强调指出，有的系统具有目的性，因为它们具有所谓的“因果力量”，而有的系统则没有目的性。但这种因果力量又从何而来呢？对此连塞尔自己都犹疑不定。他一会儿说大脑是由“特殊的东西”所构成的，一会儿又说是由别的什么东西构成的。在塞尔那里，大脑的本质似乎成了任何可以用作权宜之计的东西，它时而是区分“形式”和“内容”的不

可捉摸的要素，时而又是区分句法和语义的要素，如此等等。

在系统回答的拥护者看来，塞尔无非在说，被禁闭在房间里的那个人（以下我们称之为“塞尔的魔鬼”）只需要记忆或综合所有来自“纸上”的材料。仿佛一个人能通过任何可以设想的想象力的作用就能做到这些。写在这些“纸上”的程序代表了整个心灵以及某种对书面材料作出反应的能力特征，此能力的复杂程度可与人相媲美，因为它能够顺利地通过图林测试。有谁能够一下子“接受”对另一颗心灵的全部描写？有时候我们连背诵一段文章都会觉得很难，而塞尔硬说那个魔鬼可以记住相当于写在如果不是数百亿张也是数百万张纸上的抽象符号，而且使用时还能左右逢源，过后无需作什么更正。这些不可能存在的事情都被塞尔轻描淡写地一笔带过了，但这还不是使读者相信他的看法的主要论证的内容。相反，他的主要论证是意在掩盖数量级问题，否则有疑问的读者就会发现，差不多所有的理解力都存在于草稿纸上的亿万个符号之中，而那个魔鬼实际上却没有丝毫的理解能力。那个魔鬼是人这个事实是一个无关宏旨的枝节问题，实际上是一个会将人引入歧途的枝节问题，而塞尔却误以为这个问题至关重要。

我们这么说是完全有根有据的，不信请看塞尔本人对系统回答的赞同之辞。为此，我们首先应该将塞尔的思维实验放在一个更为广阔的背景之中。我们特别要指出，塞尔所做的实验只是许多相关的思维实验中的一个，本书的另外几章着重讨论了其他几个实验。在思维实验发生器上每个思维实验都有一个特定的“旋纽”，旋开它，可以在你的想象中产生出对人类心灵活动各种各样的模拟形象。每一个不同的思维实验都是一个“直觉泵”（丹尼特的用语），它将问题的某个侧面进行放大，促使读者得出某些结论。我们感兴趣的旋纽差不多有 5 个，当然，也

许会有人认为不止5个。

旋纽1：这个旋纽控制着构成模拟的物质的“材料”。它的装置包括：神经原和化合物；水管和水；稿纸和写在上面的符号；手纸和碎石块；数据结构和程序；等等。

旋纽2：这个旋纽控制着精确度，有了精确度才能模拟人的大脑。精确度的确定完全是任意的。它可以十分细致（原子内的粒子），也可以不那么细致，比方说细胞和突触，甚至可以更为宏观，即人工智能研究人员和认知心理学家所研究的现象：概念和观念，表象和过程。

旋纽3：这个旋纽控制着模拟的物质规模。我们这样假定，微型化的做法使我们能够将小型的水管或固体电路块网络置于一个套管中，相反，若有需要，任何化学过程也可以放大到宏观的水平上。

旋纽4：这个旋纽控制着进行模拟的那个魔鬼的体积和性质。如果这个魔鬼是一个正常大小的人，我们将称之为“塞尔的魔鬼”。如果它是个能躲在神经原中或粒子中的小精灵，我们将用豪格兰的说法称之为“豪格兰的魔鬼”，这个说法来自他对塞尔的看法的反应。这个旋纽的装置还控制着这个魔鬼是不是生物。

旋纽5：这个旋纽控制着魔鬼的工作速度。它可以使魔鬼发疯似地工作（每微秒进行几百万次运算），也可以使它慢得令人心焦（也许是每隔几秒钟进行一次运算）。

现在，通过操纵不同的旋纽装置，我们能进行各种思维实验。其中之一就是塞尔的汉语符号实验。这个实验与以下5个旋纽的关系尤为密切：

旋纽1：纸和符号

旋纽2：概念和观念

旋纽3：房间大小

旋纽4：同人一般大小的魔鬼

旋纽5：慢速装置(每隔几秒钟进行一次运算)

请注意，从原则上讲，塞尔并不反对我们假设具有这些参数的模拟能够顺利通过图林测试。他只是对这种假设蕴含的意思持有异议。

还有一个参数并不在旋纽之列，而是一种观点，是观察此实验所依据的观点。这个实验未免太单调，让我们给它点缀一下，假设那位模拟的操汉语的人是位女士，那些魔鬼(如果是生物)都是男性。现在我们可以选择从哪一个观点来观察实验，一个是魔鬼眼中的观点，另一个则是系统眼中的观点。请记住，根据我们的假设，魔鬼和模拟的女士都同样能够就是否理解并和有什么体验发表看法。然而，塞尔却一再坚持我们只能从魔鬼眼中的观点来观察这个实验。他坚持认为，不管那位模拟的女士就她的理解能力发表什么意见(当然是用汉语说的)，我们都不应该予以理睬，相反，我们应该将注意力集中于系统内部操纵符号的魔鬼。塞尔的这些话就等于说，观察实验实际上只有一个观点，而不是两个。如果有谁接受了塞尔对整个实验所作的描述，那么，他就会觉得这个说法在直觉上具有一种吸

引力，因为这个魔鬼的身材与我们相仿，口中说的是我们的语言，工作起来的速度也和我们一样快。而那位“女士”回答一个问题需要一个世纪(那也要碰得巧)，况且她写的字是“毫无意义的横直撇捺”，我们自然就很难马上与之认同了。

但是，如果我们改革某些旋纽装置，也就能轻而易举地改变观察角度。值得一提的是豪格兰的变换，他拧开了以下5个旋纽：

旋纽1：神经原和化合物
旋纽2：神经激活层次
旋纽3：大脑的体积
旋纽4：小魔鬼
旋纽5：工作极快的魔鬼

豪格兰想让我们看到这样一种情形：不幸的是，一位真正的女士的大脑是有缺陷的。她的大脑不再能够使神经介质在神经原之间互相传递。然而，不幸中的大幸是这个大脑中居住着小得和快得惊人的豪格兰魔鬼，每当一个神经原要将神经介质传到邻近的神经原上时，它都要介于其中，它刺激下一个神经原的相应的突触，其作用相当于是真正的神经介质传递过来了。这个豪格兰魔鬼动作异常迅速，他可以在亿万分之一秒的瞬间内从一个突触蹦到另一个，从来都不会落后于预定时间。因而，这位女士只要健康的话，她的大脑就会照常工作。现在，豪格兰问塞尔，这位女士还在思维吗？也就是说，她具有目的吗？或者用图林引用杰斐逊教授的话来说，她会不会只是个“人造信号”？

你也许会以为塞尔要我们听魔鬼的，与魔鬼认同，并要我

们避开主张听那位女士的，并与她认同的系统回答。但是，在对豪格兰的回答中，塞尔使我们大吃一惊，他这次竟然要我们听她的，而置魔鬼于不理，这个小魔鬼从它那小小的优势观点上咬牙切齿地诅咒着，对我们大叫大嚷道："你们这群傻瓜，别听她的！她只是个傀儡，她的一举一动全受我们的支配，全由我置放于她那些神经原之中的程序引起的。"但塞尔却没有注意这位豪格兰魔鬼的警告。他说："她的神经原依然具有合适的因果力量；它们只是需要从魔鬼那里获得一臂之力，仅此而已。"

我们可以在塞尔原来的装置与这个经过修改的装置之间作一些比较。与原来的"稿纸"相对应的是这位女士的大脑突触。与原来写在"纸"上的人工智能程序相对应的是这位女士大脑的整个构造，它等于一个庞大的指令，指示魔鬼在什么时候、知道怎样去刺激哪些突触。与原来在纸上写"没有意义的横直撇捺汉字"活动相对应的是刺激她的突触的活动。假设我就沿用现在这样的装置，只是将体积和速度旋纽作了一些变动，那么我们将把这位女士的大脑放得和地球一样大，那个魔鬼也就放大成了"与我们一般大小"的塞尔魔鬼而不是豪格兰魔鬼，假设这位塞尔魔鬼的行动速度在人类看来是很合理的，而不是在一微秒之内就已在这个形同地球的大脑里走过了千万里，那么，塞尔现在让我们与哪个魔鬼认同呢？我们不想枉作猜测。但是我们认为，如果系统回答在前面已使人折服，那么，在此它也应该令人无可非议。

必须承认，塞尔的思维实验极为生动地提出了这个问题：理解一种语言究竟是什么意思？我们想扯开去对这个问题稍加议论。请考虑一下这个问题："究竟什么样的处理某种语言的书面或口头符号的能力才算得上是对这一语言的真正理解？"鹦鹉

能学说英语，但不懂英语。电话报时台的那个女性的声音能够准确地报出时间，但它并不是一个理解英语的系统的喉舌。在那个声音的背后没有心理活动，它的心理基质被撤去了，只剩下貌似人类的东西。也许一个天真的儿童会不明白，怎么会有人干这样单调的工作，而且干得这样丝毫不出差错。我们会觉得这种想法很好笑。当然，如果那个声音的背后是一个能够顺利地通过图林测试的人工智能程序，那么，此事又该当别论了！

假定你自己眼下正在中国讲学。你知道自己在用英语进行思维，然后，在最后一分钟里(事实上是在最后的一刹那里)，你使用了转换规则，这些规则以一种奇特的、“毫无意义”的方式将英语思维变成了控制你嘴部和声带的指令，然而，你的中国学生却听得津津有味。当他们举手提问时，他们说的汉语你虽一窍不通，但这不要紧，你胸有成竹，因为你可以迅速地使用某些反转规则，把这些话的英语意思揭示出来……如果是这样，你觉得你真在说汉语吗？你觉得你对操汉语的人的心理有所领悟吗？或者说，你是否真能想象这种情景？它现实吗？用这种方法真能说好外语吗？要学会汉语，一般的方法是“必须学习用汉语进行思维”。但这话又是什么意思？学过外语的人都会有这样的体验：外语的声音很快就变得“听不见”了，你不是听见了声音，而是凭借了声音，正如临窗远眺时，你不是看见了窗户，而是凭借窗户看见了景色。当然，如果你费尽心思，也可以使一门你十分熟悉的语言光闻其声而不明其意，就像如果你愿意的话，你也尽可以望着窗玻璃出神；但你不能两者兼得，也就是说当听到一种语言时，你不能又觉得懂又觉得不懂。因此，一般来说，人们所听到的是语言的意义。如果有的人学习一门语言纯粹是为了听它美妙的语音，他未免会觉得有点失望。然而一旦把握了这些声音，即使你不再是天真地听它们，

也是一种妙不可言的体验(如果你将同一种分析运用到音乐欣赏中去，那将十分有意思，尽管人们对于光听音乐的声音与听其中的“意义”之间的差别知之甚少，但这个差别看来确实存在)。

学习外语意味着超越自己的母语，意味着将新学的语言正确地融入思维所采用的母语。用新学的语言进行思维与用母语进行思维必须能同样轻松(至少是差不多轻松)。一门新语言的习惯如何一层一层地深入直到最终植入大脑神经原的过程，还是个未解之谜。但是，有一点是肯定的，掌握一门语言决不是让你的“英语系统”去贯彻一套程序规则，从而能将一门语言当作一套没有意义的声音和符号来对待。这门新语言一定通过某种途径与你的内部表象系统，即概念、意象等的总和融为一体，就像你的母语与这个表象系统融为一体一样。为了更精确地考虑这个问题，你必须对执行层次有一个十分明确的概念，在计算机科学中，这个概念作用很大。

一个系统能够“仿真”另一个系统，对此观念计算机科学家们已习以为常。事实上，这个观念源于一条早在1936年已被艾伦·图林证明的公理：任何一台万能数字计算机都能够模拟其他的万能数字计算机，在外界看来唯一不同在于速度的差异。“仿真”一词专指计算机模拟计算机，而“模拟”一词则指计算机模仿其他现象，如飓风、人口、曲线、大选乃至计算机的使用者。

模拟与仿真之间的主要差别在于，模拟总是无限接近于被模拟的对象，它取决于有关现象的模型本质；而仿真则强调精确得丝毫不差。仿真的逼真度极高，比方说，一台Σ—5型计算机仿真了另一台不同结构的计算机，比如一台DECPDP—10型计算机，可以使操作者丝毫觉察不出。将一种结构置于另一种结

构之中的做法形成了所谓的“虚拟计算机”，在上面的例子中就是一台虚拟DEC—10型机。在虚拟机的背后总隐藏着另外一台计算机。它可能是一台同样型号的机器，也可能是另一台虚拟机。安德鲁·坦南鲍姆在《结构计算机组织》一书中用虚拟机这一概念来解释如何将大型计算机系统视作虚拟机相叠的结果，当然，最底下的那台机器是一台真正的计算机！但是在任何情况下，各个层次都被严密地隔离开来，就像塞尔的魔鬼无论如何也无法与操汉语的人沟通，尽管他就是这个人的一部分(塞尔的魔鬼不谙汉语，如果我们假定有一位翻译在场，那么，他们会谈些什么呢？这个问题非常有意思)。

从理论上讲，两个不同的层次有可能互为沟通，但一般说来，人们不赞成这么做，层次间的混淆是不允许的。然而，人类“系统”在学外语时发生的正是这种情形——抹煞两个执行层次之间的差别。外语并不会像软体寄生物那样寄生在母语的顶上，实际上，它也像母语一样(或者说基本上一样)深深地根植于硬件之中。学习外语就意味着深刻地改变一个人的内在“机器”，神经原激活的方式也发生了巨大而连贯的变化，这一连串的变化，为更多层次的实体(即符号)创造了互相引发的新方式。

在计算机系统中，情形也一样。一个较高层次的程序必须有办法在执行这个程序的“魔鬼”体内引起变化。这种情况对于计算机科学目前的做法来说是生疏的，因为目前我们在叠加层次的过程中对各层次加以严格的隔离。较高层次既能作环状运动又能影响较低的层次，这种能力具有魔幻的色彩，我们觉得它与意识的核心颇为接近。也许有一天我们能够证明它在推动设计更为灵活的计算机方面，当然，也在加深认识人工智能方面起着极为重要的作用。尤其是，如果我们要对什么是“理解”这个问

题作出令人满意的回答，那就无疑应该更清晰地描述出符号处理系统内的各个层次如何相互依赖并相互影响的。总而言之，这些概念已经证明是令人难以捉摸的。如果对它们能有清楚的理解，那将大有益处。

在这场关于多层次的颇为混淆的讨论中，你也许会开始怀疑，这里所说的"层次"究竟是什么意思。这是一个非常难的问题。只要各层次互为隔离，像塞尔的魔鬼与操汉语的女士那样互不接触，这个问题是十分明显的。但是一旦它们开始混同，那么，请你注意：塞尔也许会承认在他的思维实验中有两个层次，但是，他不愿意承认存在两个观点，也就是说存在两个真正的生命，这两个生命有感觉并"具有经验"。一旦我们承认某些计算机系统也具有经验，塞尔顿时就会忧心忡忡，因为在他看起来这就像潘多拉的盒子，一旦打开，骤然间就会"心灵遍野"，从肠胃、肝脏乃至汽车发动机里都会有心灵。

如果我们竭力想将系统描写成人工智能程序的例示，那么，任何系统都能够被说成具有信仰、感觉等等。看来塞尔相信这个说法。很明显，这个说法令人不安，因为它将导致泛灵论。事实上，塞尔坚信人工智能的研究者已经不知不觉地信奉了泛灵论的世界观。

塞尔为了摆脱自己布下的陷阱而坚持认为，你在无机物中看见的所有那些"信仰"和"感情"，以及你在四处都见到的那些心灵并不是真的，而是"假的"。它们没有目的性，它们没有大脑的因果力量！（当然，塞尔会提醒大家别将这些概念与"灵魂"这个朴素二元论的概念混为一谈。）

我们摆脱这个陷阱的办法则是坚持认为，这个陷阱是不存在的。心灵遍野的观点是不正确的。我们认为：就像大脑不可能在汽车发动机或肝脏中存在那样，心灵也不可能出现在汽车

发动机或肝脏之中。

我们觉得有必要在这一点上多谈几句。如果你能够在蠕动着的肠胃中看到复杂的思维过程的话，那么，难道不能从汽水的泛泡型式中译解出肖邦E小调钢琴协奏曲，从瑞士奶酪的小孔中译解出整部美国历史吗？当然可以，不仅能用英语，还能用汉语来译解。世上万物毕竟是都有其含义的。你可以用《哈姆雷特》的结构来译解巴赫的第2号勃兰登堡协奏曲，而《哈姆雷特》的结构则可以从你狼吞虎咽下去的最后一块生日蛋糕的结构中译解出来(如果你知道译码的话)。

在所有这些例子中，问题就在于当你事先还不知道要认识什么时，就去分析译码了。否则，你就可以根据一个任意构造的后天的译码从一场足球赛或一片草叶中得出对任何人的心理活动的描述。但这决不是科学。

当然，心灵的复杂程度不尽相同，但是，心灵之所以是心灵正是因为它具有一个复杂的表象系统，而在汽车发动机或者肝脏中却找不到这样一个不断自我更新的表象系统。当然，正如人们可以认为埃及金字塔或英格兰巨石群、巴赫的音乐、莎士比亚的戏剧会有别的额外的意思一样，你也可以认为轰鸣的汽车发动机有着心灵的意思，换句话说，你可以用牵强附会的方法来随时满足释解者的各种愿望。但是，我们想这不会是塞尔的本意吧(尽管我们假定这就是他的本意)。

心灵存在于大脑之中，也可以存在于程序计算机之中。如果有一天这样的机器被造了出来，它们的因果力量不会来自于构成它们的物质本身，而是来自于它们的设计以及在其自身运行的程序。而如果我们想知道它们是否有因果力量，就得与它们交谈，倾听它们想说的话。